2011-06

An exploration of the communications environment within the System of Systems Survivability Simulation (S4)

Edmiston, Russell J.

Monterey, California. Naval Postgraduate School

http://hdl.handle.net/10945/5647

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

AN EXPLORATION OF THE COMMUNICATIONS ENVIRONMENT WITHIN THE SYSTEM OF SYSTEMS SURVIVABILITY SIMULATION (S4)

by

Russell J. Edmiston

June 2011

Thesis Advisor: Thomas W. Lucas
Second Reader: Dashi Singham

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** 7-6-2011	**3. REPORT TYPE AND DATES COVERED** Master's Thesis 2009-07-01–2011-06-30	
4. TITLE AND SUBTITLE An Exploration of the Communications Environment within the System of Systems Survivability Simulation (S4)			**5. FUNDING NUMBERS**
6. AUTHOR(S) Russell J. Edmiston			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) Army Research Laboratory, Adelphi, Maryland			**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number XXXX.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited			**12b. DISTRIBUTION CODE**

13. ABSTRACT (maximum 200 words)

The U.S. Army is transforming into a network-centric modular force. Developed by the Army Research Laboratory's Survivability/Lethality Analysis Directorate, in conjunction with New Mexico State University's Physical Science Laboratory, the System of Systems Survivability Simulation (S4) is a detailed, agent-based, Monte Carlo simulation designed to conduct survivability, lethality, and vulnerability assessments of military platforms connected via communications networks to other platforms on the battlefield. This thesis explores key parameters that make up S4's communications environment, using design of experiments and data farming tools to determine if any or all are having unintended interactions. This thesis concludes that the explored parameters generally perform as intended, decreasing or increasing communication performance as the environment becomes respectively more or less restrictive. However, the level of influence of the parameters varies greatly, questioning either the level of realism to how the parameters are modeled or their necessity to the simulation. In addition, the variation across input settings and replications demonstrates the value of being able to efficiently explore multiple factors and take many replications. As a pilot study, these results and methodology pave the way for enhanced analytical capability with S4 and its continued verification, validation, and accreditation.

14. SUBJECT TERMS			**15. NUMBER OF PAGES**
System of Systems, S4 System of Systems Survivability Simulation, Design of Experiments, Nearly Orthogonal Nearly Balanced Mixed Design, Survivability Lethality and Vulnerability Assessment, Communications Environment			83
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

AN EXPLORATION OF THE COMMUNICATIONS ENVIRONMENT WITHIN THE SYSTEM OF SYSTEMS SURVIVABILITY SIMULATION (S4)

Russell J. Edmiston
Major, United States Army
B.S., Arizona State University, 2000

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
June 2011

Author: Russell J. Edmiston

Approved by: Thomas W. Lucas
Thesis Advisor

Dashi Singham
Second Reader

Robert F. Dell
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The U.S. Army is transforming into a network-centric modular force. Developed by the Army Research Laboratory's Survivability/Lethality Analysis Directorate, in conjunction with New Mexico State University's Physical Science Laboratory, the System of Systems Survivability Simulation (S4) is a detailed, agent-based, Monte Carlo simulation designed to conduct survivability, lethality, and vulnerability assessments of military platforms connected via communications networks to other platforms on the battlefield. This thesis explores key parameters that make up S4's communications environment, using design of experiments and data farming tools to determine if any or all are having unintended interactions. This thesis concludes that the explored parameters generally perform as intended, decreasing or increasing communication performance as the environment becomes respectively more or less restrictive. However, the level of influence of the parameters varies greatly, questioning either the level of realism to how the parameters are modeled or their necessity to the simulation. In addition, the variation across input settings and replications demonstrates the value of being able to efficiently explore multiple factors and take many replications. As a pilot study, these results and methodology pave the way for enhanced analytical capability with S4 and its continued verification, validation, and accreditation.

THIS PAGE INTENTIONALLY LEFT BLANK

Table of Contents

List of Figures

List of Tables

THIS PAGE INTENTIONALLY LEFT BLANK

List of Acronyms and Abbreviations

AA	Assembly Area
AOI	Area of Interest
ANOVA	Analysis of Variance
ARL-SLAD	Army Research Laboratories-Survivability/Lethality Analysis Directorate
BIC	Bayesian Information Criterion
BN	Battalion
bps	Bits Per Second
C2	Command and Control
CDR	Commander
CO	Company
COA	Course of Action
CSV	Comma Separated Value
DOE	Design of Experiments
DMP	Decision-Making Process
DVD	Digital Video Disc
FBCB2	Force XXI Battle Command Brigade and Below
GB	Gigabyte
GUI	Graphical User Interface
HMMWV	High Mobile Multipurpose Wheeled Vehicle
km	Kilometer
LDR	Leader
LOESS	Locally Weighted Scatterplot Smoothing
MB	Megabyte
METT-TC	Mission, Equipment, Terrain, Troop, - Time, Civil
MIP	Mission Information Packet
MTU	Maximum Transmission Unit
NATO	North Atlantic Treaty Organization
Net	Network
NMSU-PSL	New Mexico State University-Physical Science Laboratory
NOLH	Nearly Orthogonal Latin Hypercube
NONBMD	Nearly Orthogonal Nearly Balanced Mixed Design

OBJ	Objective
OLH	Orthogonal Latin Hypercube
PLT	Platoon
PCOM	Probability of Communications
Recon	Reconnoiter or Reconnaissance
S4	System of Systems Survivability Simulation
SA	Situational Awareness
SINCGARS	Single Channel Ground and Airbourne Radio System
SITREP	Situation Report
SLVA	Survivability, Lethality, and Vulnerability Assessment
SoS	System of Systems
TCP	Transmission Control Protocol
UAV	Unmanned Aerial Vehicle
VV&A	Verification, Validation, and Accreditation
XML	Extensible Markup Language

Executive Summary

The purpose of this thesis is to conduct quantitative analysis of the communications environment within the System of Systems Survivability Simulation (S4). Additionally, a methodology is provided that dramatically enhances the information that can be obtained from S4. This research also provides a first step toward verification, validation, and accreditation (VV&A) of the model, which will then enable its use by the Department of Defense to conduct survivability, lethality, and vulnerability analysis. This summary provides an explanation of the need for such a model, the justification of the methodology, an overview of the analysis conducted, and the conclusions and recommendations for the future exploration and development of S4.

The U.S. Army has been transforming into a modular force, with a focus on interconnecting combat systems through a vast communications network designed to improve Command and Control (C2), the distribution of information, and consequently, survivability (Davidson, Pogel, & Smith, 2008). As a result, the success of one military system is dependent on the success of the other systems it is connected to on the battlefield. This concept of a System of Systems (SoS) requires that each system be measured not by its solitary performance, but rather its contribution to the larger system (Bernstein Jr., Flores, & Starks, 2006). A full-scale test and evaluation of an SoS would come at great cost and effort, and with numerous limitations. A simulation, however, would be more cost effective and allow for greater flexibility and multiple experiments. The Army has a history of using models and simulation to conduct many different types of analysis and there now exists an urgent need for a simulation that models the effects of a network-centric force (Zacharias, MacMillan, & Hemel, 2008). S4 was developed to fill that need. Created in cooperation between the Army Research Laboratory's Survivability/Lethality Analysis Directorate (ARL-SLAD) and New Mexico State University's Physical Science Laboratory (NMSU-PSL), S4 is a detailed, agent-based, Monte Carlo simulation designed to conduct survivability, lethality, and vulnerability assessments (SLVA) of military platforms within an SoS.

With communications being the foundation of a network-centric force, this thesis focuses on key parameters chosen by NMSU-PSL that make up a critical portion of the communications environment. In addition to communications, S4 models terrain, maneuver, sensing, ballistic engagements, perceptions, and decision making. Shown in Figure 1, these processes are layered with communications as the conduit between what actions/events an agent encounters in

the simulated world and its perceptions and decision-making ability. Communication is similarly layered by the equipment modeled, such as radios and antennas, and the environment they operate in, comprised of effects from both the physical environment and imposed procedures and protocols. This thesis analyzes 11 parameters that make up this communications environment; however, two of the parameters are flags that enable their partnered parameter to be varied. The remaining nine parameters include: *Antenna Gain*, *Link Refresh—Distance*, *Link Refresh—Time*, *Max Transmission Control Protocol (TCP) retries*, *TCP retransmit interval*, *Bits in Data Unit*, *Bits of Overhead*, *Available Bandwidth*, and probability of communication success *(PCOM)*. These parameters affect either data communications or both data and voice communications, and contain continuous, discrete, and categorical values. To evaluate them, a scenario was developed that minimized all extraneous processes, such as ballistic engagements, while increasing the number of communications in order to stress the modeled networks.

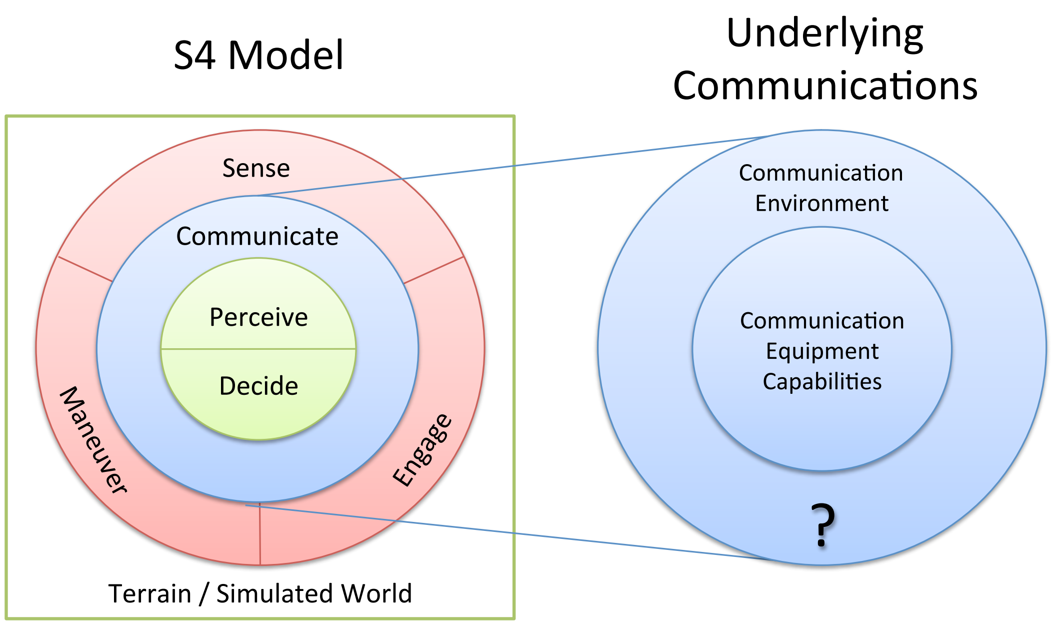


Figure 1: At the right, the actions and processes modeled by S4 are pictured as concentric circles (After Bernstein Jr. et al., 2006). At left is the interpretation for this thesis how the communications environment is similarly layered.

This thesis utilized the data farming process as described by Horne and Meyer (2004), incorporating a design of experiments (DOE) to efficiently explore an entire landscape of outcomes. However, the process of developing a DOE faced three main challenges: (1) a limited computing budget; (2) the need to handle 11 continuous, discrete, and categorical values as well as restrictive parameter rules; and (3) a random number generator that feeds multiple processes within the simulation. NMSU-PSL chose to conduct the simulation runs in-house, which prevented the use of cluster computing. As a result, the number of experiment runs—and therefore design points—had to be low enough that they could be completed on one computer over a reasonable amount of time. This was achieved by building a custom design using a newly developed Nearly Orthogonal Nearly Balanced Mixed Design (NONBMD) (see Vieira & Sanchez [2010]).

The stochastic nature of S4 dictates the need for replication of each design point. The simulation, however, uses only one random number stream for multiple probability processes without a clear understanding of when or in what order each is drawn. Differences between the results of the simulation, therefore, cannot be guaranteed to be only a result of changes to parameter values and independent errors. Using a sample of 12 runs, the variance caused by changing only the random seeds was calculated. From this value, it was determined that in order to create a tight confidence interval around the results, each design point would need to be replicated 15 to 25 times. However, although the number of design points created by the NONBMD was low enough to allow for replication, it was still too big to allow for more than five replications due to the restrictive run budget. The final number of runs conducted was 2,160 (144 design points $*$ 5 replications $*$ 3 scenario variants).

Two responses were used to analyze the output from the experiment. The primary response for the analysis is Total Communication failures. The number of failures were also extracted by type of failure (of which there are four) and by agent (of which there are 21). The second response—Operational Delay—is intended to show the operational consequences of a degraded communications environment. S4 produces a large amount of output, such that not all of it could be transferred from NMSU-PSL to the Naval Postgraduate School (NPS). As a result, Operational Delay was calculated indirectly and under the assumption that changes to the communications environment do not affect agent mobility. Using the same 12 runs used to determine the variance caused by changing the random seeds, average agent travel times were determined. Those times were then used, along with extracted data, to calculate the changes in time between the reporting of a critical event and the action that event is designed to trigger.

Analysis of the output included observations from response distributions, effects caused by changes to the scenario, impacts on different agents, relationship between the two responses, and meta-models (regression and partition trees) in order to determine parameter influence. The results indicate that there is tremendous variability across design points. Thus, achieving robust solutions requires that many factors be explored. The results from this analysis also showed that the level of influence varied greatly among the analyzed parameters. Although *PCOM* and *Max TCP Retries* were consistently shown to be the top two influential parameters, the order of influence of the others changed based on the response and meta-model. Some parameters were shown only to be influential as an interaction term and others only under certain ranges. Regardless, the simulation performed as expected. When the parameter values created a poor communications environment, both the number of failures and Operational Delay increased. The inverse was also shown to be true: when the parameter values created a good environment, both responses had improved values. Unfortunately, a design flaw—changing the parameter values for all agents in the simulation—added to the variability, and created noise that limited the use of the Operational Delay response to determine the sensitivity of S4's modeled decision making.

This thesis concludes that the parameters of most influence in this scenario are *PCOM* and *Max TCP Retries*. Due to the varying level of influence of the other parameters, it is recommended that they be checked for either their modeled accuracy with the real world or their necessity to the simulation. Most importantly, however, based on the realization that the simulation is generally performing as expected, it is determined that the communications environment is indeed playing a logical role in the outcome of the simulation, as well as being a conduit between the communications equipment and the equipment users. Due to the large number of parameters within S4, as well as the multiple processes modeled, it is recommended that further studies of S4 incorporate the analysis, methodology, and lessons learned from this thesis.

Acknowledgements

This thesis is more of a collective effort than a solitary one. My advisor, Dr. Tom Lucas, and my second reader, Dr. Dashi Singham, provided tremendous leadership, support, and mentorship throughout this entire process. Their guidance and motivation not only played a large role in making this thesis a product worth reading, but made me a better analyst.

I would also like to thank the great support from the S4 staff at NMSU-PSL, in particular Dr. Jayashree Harikumar and Dr. Roger Hartley. PSL put a huge effort into this thesis, from building the scenario to conducting all the simulation runs. Moreover, their responsiveness to, and patience with, my questions and issues was outstanding. With such a quality staff guiding its development, S4 is destined to become an extraordinary analytical tool.

Finally, I would love to thank my wife, Emma, for her unwavering support and patience throughout the development of this thesis, as well as during the completion of the Operations Analysis curriculum at NPS. Without her, I would truly not be where I am today.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 1: INTRODUCTION

For the past several years, the U.S. Army has been transforming into a modular force in order to better conduct asymmetric operations across the full spectrum of conflict. According to the Army's Transformation Roadmap, a foundation of this new force is the use of modern combat systems interconnected to each other across a vast communications network designed to increase survivability through better Command and Control (C2) and distribution of critical information (Army Transformation Office, 2004). However, there are many challenges with putting together a network-centric force:

- Effectively integrating technology with both new and old platforms.
- Defending the network itself against a less predictable and more innovative enemy.
- Harnessing technological advances that outpace the Army's procurement process with consideration of an enemy that has no bureaucratic procurement process and is capable of rapidly fielding, and putting to deadly use, new equipment, technology, and means of attack (Davidson, Pogel, & Smith, 2008; Bernstein Jr., Flores, & Starks, 2006).

To address these challenges, the Army is in need of an analytical tool that, without incurring additional procurement time, can quantify and evaluate how one system performs within, or adds value to, the greater system that is comprised of the interconnected systems across the battlefield.

One model that seeks to do just that is the System of Systems Survivability Simulation (S4) model. S4's commissioned purpose is to conduct survivability, lethality, and vulnerability assessments (SLVA) specific to System of Systems (SoS)-level effects related to threat, computer network operations, electronic warfare, and ballistics events (Smith, Bernstein Jr., Hartley, & Harikumar, 2010). Modeling communications networks, military decision making, and military equipment using actual equipment performance data, S4 has shown great promise in accomplishing that purpose. However, to date, the model has not been taken through a verification, validation, and accreditation (VV&A) process. VV&A is a modeling prerequisite that helps to ensure that flawed or biased models do not become mainstream tools. Without VV&A, it is unknown as to whether or not a model's produced results are accurate or correctly interpreted. This study analyzes S4 using design of experiments (DOE) and data analysis to begin the process of quantitatively measuring the limitations and constraints of the model.

1.1 BACKGROUND AND LITERATURE REVIEW

Traditionally, the Army evaluates new equipment by comparing the stand-alone capabilities of the new system against the stand-alone capabilities of the existing system. A network-centric force, however, demands much more than these simple measures (Starks & Flores, 2004). With systems tethered to work together with other systems to enable combat power and information superiority, the success of a solitary system is not independent, but rather dependent on the success of every system on the battlefield with the network itself serving as a critical component (Davidson et al., 2008; Starks & Flores, 2004). Each system, therefore, must be measured not by its solitary performance, but how well it contributes to the larger system—or as an SoS (Bernstein Jr. et al., 2006).

> System of Systems (SoS): A collection of interlinked and mutually dependent systems that has properties and capabilities well beyond the simple union of the independent attributes of its constituent systems. (Smith et al., 2010, p. 9)

To conduct a full-scale test and evaluation at the SoS level would require tremendous time, effort, and cost. However, a simulation that could capture the dependent interactions of SoS would not only reduce the effort required, but it would also provide the evaluators and decision makers with the flexibility and insight of multiple runs using different scenarios, equipment, organizational configurations, or environments. The Army has a long history of using models and simulation to conduct analysis for planning forecasts, training rehearsals, as well as design and evaluation for acquisition (Zacharias, MacMillan, & Hemel, 2008). Furthermore, there is an "urgent need" for research regarding network-centric operations (Committee on Modeling & Simulation for Defense Transformation, 2006). However, the challenge of producing such a simulation that "specifically account[s] for information technologies; support of battle command, along with the traditional engagement-based warfighting functions, such as move and maneuver and fires" (Davidson et al., 2008, p. 155) is just now being explored.

S4 was created in cooperation between the Army Research Laboratory's Survivability/Lethality Analysis Directorate (ARL-SLAD) and New Mexico State University's Physical Science Laboratory (NMSU-PSL). ARL-SLAD conducts, among other things, modeling and simulations of military systems and provides analysis of those experiments to senior leaders and developers. They conduct this mission in order to ensure that military personnel and equipment survive and function effectively in hostile environments. NMSU-PSL has a long history of aiding military test and evaluation, dating back to 1946. Since 2004, the NMSU-PSL/ARL-SLAD team, comprised of experts in, among other areas, computer modeling, communication operations, and ballistics effects, has worked diligently to produce a model that analyzes the survivability performance of military platforms and equipment, both small and large, in conjunction with other platforms and equipment on the battlefield.

As S4 continues to develop and improve as an SoS SLVA tool, one of the steps it must complete to be used by the Department of Defense is a VV&A process (Department of Defense, 2009). To verify the model, the developers must show that the produced output accurately reflects its desired purpose. Currently, there is no clear understanding as to to how some of the parameters in the model interact with each other or their value to the model. Since the communications network is the backbone to network-centric warfare, the best place to begin verification is with S4's modeled communications environment.

1.2 RESEARCH QUESTIONS

The intent of this thesis is not to conduct a complete VV&A process, but rather to conduct a quantitative analysis of the S4 model. To that end, this thesis is guided by the following questions:

1. Which communications factors have the most influence on the model's output?
2. How sensitive are the model's decision-making processes to the success or failure of communications?
3. Is the model's output a result of the equipment capabilities being analyzed or a product of how the communications environment is modeled in S4?

1.3 BENEFITS OF THESIS

This thesis provides a way ahead for the continued development and VV&A of S4. Regardless of whether or not adverse interactions are identified, the provided methodology can be used to examine additional parameters within the simulation. Furthermore, this process highlights model inadequacies not previously identified or explored by PSL. The results and conclusions of this thesis can be used to explore and implement changes to the model, thus strengthening the use of S4.

1.4 METHODOLOGY

This thesis analyzes 11 communications parameters that have been selected by analysts at NMSU-PSL and make up a significant part of the communications environment within S4. This is accomplished through an iterative process that includes a DOE utilizing nearly-balanced, nearly-orthogonal parameter selection that stresses the thresholds of those parameters, exploration of the model, and finally, analysis of the output, in order to examine the parameter interactions within the model and determine if any or all are providing unintended results or behaviors. This is accomplished using one simple scenario with three variants, each of which place heavy emphasis on communications over movement, engagements, or other battlefield actions.

CHAPTER 2:
S4 OVERVIEW

In order to provide SLVA, at its core, a model must answer the question: "what is the effect of the performance of a given piece of equipment on the ability of a combat unit to accomplish its objective?" (Hartley, 2010, p. 5). Information needed to answer this question includes:

1. Specifications of all the equipment and platforms in use.
2. Interactions between types of equipment.
3. Interactions between equipment and the surrounding environment.
4. Decisions, and distribution of those decision, that are made to employ the equipment.

S4 provides all of this information and more. It is a complex model, with numerous inputs generating detailed outputs that is undergoing continuous technological advances and improved methodologies. The following paragraphs attempt to condense the growing literature on S4 into generalized concepts detailed enough for the purpose of understanding this thesis.

2.1 DESIGN AND METHODOLOGIES

S4 is a collection of integrated software components that models, in increasing levels of fidelity, sensing, terrain, movement, decision making, ballistic engagements, and communications through the use of an agent-based, partly Monte Carlo simulation (see Figure 2.1). The agents within S4 represent dismounts, platforms, key leaders, or leaders on platforms, and their actions follow a "sense, decide, and action loop" (Bernstein Jr. et al., 2006, p. 7). In each time step, an agent senses what it knows and perceives about the current situation, either from what it learned itself or through communication with other agents. Based on that information, the agent evaluates the situation and decides what actions are to be taken, if any at all. Lastly, the agent attempts to carry out the actions it decided on. This process continues for each agent until the simulation is terminated as a result of a given stopping condition. The time and location of each sense, decision, and action is recorded allowing for both playback and detailed analysis of each agent or process (Bernstein Jr. et al., 2006).

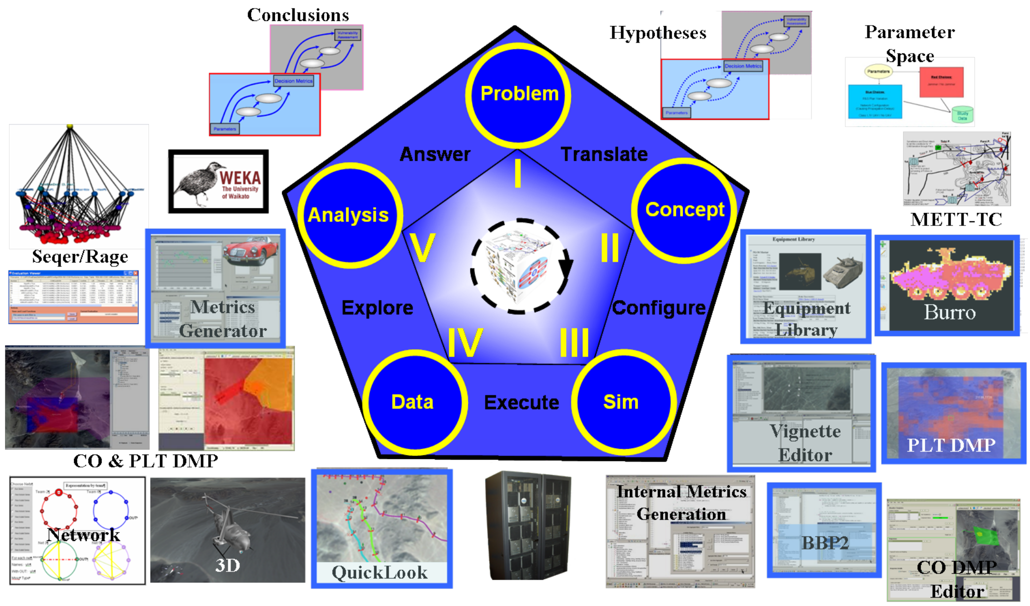


Figure 2.1: Graphical depiction of the inputs, outputs, processes, and tools used by S4 (From Smith et al., 2010).

2.1.1 Environment and Equipment

To conduct their analysis, ARL-SLAD uses equipment data pertaining to manufacturer specifications and performance results from live field tests. With that same data, NMSU-PSL utilizes simulation technologies to model equipment within S4 to a high degree of precision. Depending on the type, equipment in S4 contains information regarding physical dimensions, speeds and rates, ranges, weapon capabilities, armor performance, ballistic impacts, damage assessments, and much more (Bernstein Jr. et al., 2006). The environment within S4 is three-dimensional terrain. Platforms traverse this environment using the North Atlantic Treaty Organization (NATO) Reference Mobility Model incorporated into S4 to address vehicle dynamics and obstacles (Smith et al., 2010; Vong et al., 1999). As a result, the model provides plenty of context for agents to make informed decisions and the results of the simulation to be relevant to reality (Bernstein Jr. et al., 2006, p. 6).

2.1.2 Decision Making

The decision-making process (DMP) within S4 is explicit. Based on its position in the command hierarchy, an agent is assigned a specific DMP role. Each DMP is given specific rules to follow, such as objectives to meet, criteria on how to evaluate known and perceived information, how to prioritize targets, and contingency plans to take if a given objective is threatened or cannot be completed. There are currently 17 different DMPs modeled within S4 that dictate different aspects of the simulation, whether that be movement and maneuver, C2, intelligence, or fires. Since the tasks and purposes of each DMP differ greatly, S4 uses multiple techniques to model them (Hartley, 2010).

The complexity and number of decisions made by a DMP has an approximately inverse relationship to that agent's placement within the command hierarchy. On the top of the hierarchy, Battalion and Company DMPs make generalized decisions utilizing a "library of situation-response" templates (Davidson et al., 2008, p. 157). Based on their knowledge and perceptions of locations and activities of friendly and enemy forces, these command DMPs evaluate Courses of Action (COA) templates and select the best fit. The chosen COA is then issued to the subordinate agents through a Mission Information Packet (MIP). Although the situation is constantly changing, it rarely changes enough to force the command DMP to devalue its current best-fit COA for another. On the low end of the hierarchy is the Platform DMP, which is required to make constant decisions regarding its movement, threat assessment, weapon selection, etc. The most advanced DMP in S4 is considered to be the Platoon Leader DMP, which among other things, projects future states to select tasks, translates received mission orders into assignments for the members of the platoon, and issues both speed and target information (Davidson & Pogel, 2010).

A key takeaway from S4's modeled DMPs is that they are at the core of the simulation's sense, decide, and action loop. No movement occurs in the simulation until a decision has been made. Even at the start of a simulation, agents must first report their locations and statuses up their respective chains of command until the applicable DMP is reached, which then evaluates that information and issues the first mission order. However, since those reports are sent via communication networks, the DMPs within S4 are dependent on the success or failure of communications.

2.1.3 Output and Analytic Tools

S4 generates enormous amounts of output pertaining to current agent states, communications, decisions, perceptions, ballistic events, electronic warfare, and much more. Produced primarily in extensible markup language (XML) or comma-separated value (CSV) formats, these output files are typically separated by agent or process. There also exists a "Playback" file that contains all movement, sensing, and engagement information by each agent, frame-by-frame. This level of detail enables NMSU-PSL analysts to answer just about every possible question related to SLVA. Simple metrics can be analyzed, such as network performance, graphically charting an agent's perception against the "ground truth" of the simulation's reality, or even the time of events that drove key decisions to be be made. More detailed visual analysis can be made using NMSU-PSL's *QuickLook* "data contextualizer." This enables the analyst to reconstruct the simulation at each time-step and cartographically depict specific events such as ballistic engagements (see Figure 2.2). Although *QuickLook* provides a terrific picture of these events, it only shows the results from a single run, rather than a statistic and confidence interval from multiple runs. Since S4 uses a Monte Carlo process, one run is but one realization of a random process.

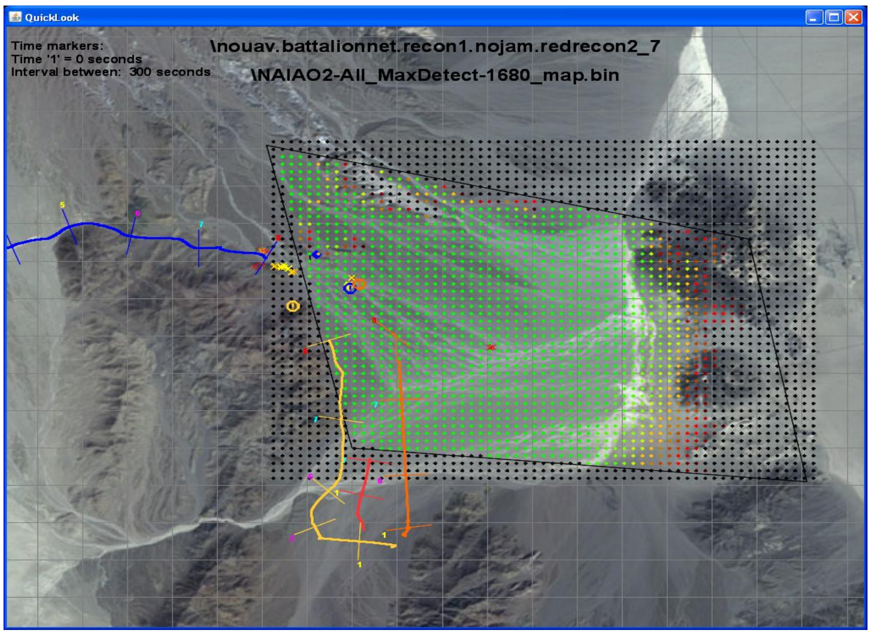


Figure 2.2: Image from a *QuickLook* screen showing agent activities by time-step (From Smith et al., 2010).

2.1.4 Random Number Generation

Many of the inputs to S4 are probabilities that an event will occur. These probabilities include detection, ballistic impacts, and communication success. Instead of using these probabilities in a deterministic sense, S4 implements a stochastic process (i.e., random number draws) that introduces randomness to these interactions between agents, the simulation environment, as well as within the DMP (Hartley, 2010). When an action arises within the simulation that has a probability associated with it, a random number is generated and used to determine the success or failure of that action. To accomplish this, S4 uses one initial random seed to generate one stream of random numbers for the entire simulation. Each event probability then draws from that single stream when adjudication is required.

2.2 FOCUS OF THESIS

This thesis begins the exploration of S4 by starting with the main ingredient to network-centric warfare: the network, or more simply, the communication between units on the battlefield. This study argues that the communications within S4 can be split into two categories: equipment capabilities and environment. In the equipment capabilities category, S4 models voice and data communications as either generic equipment effects or equipment specific to the Army's inventory, such as Single Channel Ground and Airborne Radio Systems (SINCGARS), Force XXI Battle Command Brigade and Below systems (FBCB2), and Mast Antennas. In the environment category, S4 models both the physical environment, such as propagation of radio waves, and the procedural or protocol environment, such as the number of times an agent attempts to resend a failed message. Figure 2.3 attempts to display this concept graphically. Since it is the environment that restricts or enhances the operation of the equipment, equipment performance depends on the environment it operates in. Equipment can be modeled by manufacturer specifications and results from performance tests. Environmental effects, however, can be a little more difficult to get right. For this reason, this thesis focuses on key environment parameters and their effects within the model.

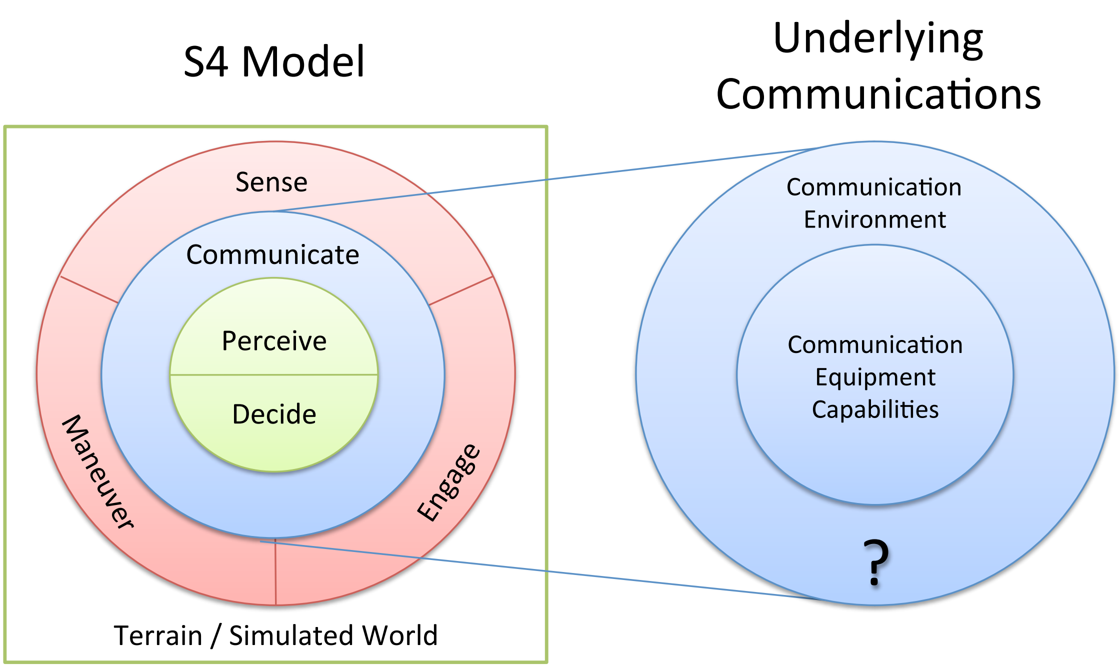


Figure 2.3: At the right, the actions and processes modeled by S4 are pictured as concentric circles (After Bernstein Jr. et al., 2006). At left is the interpretation for this thesis of how the communications environment is similarly layered.

Listed in Table 2.1, the parameters of interest for this thesis do not make up all of the environmental parameters. However, they were chosen specifically by NMSU-PSL due to their great significance to the simulation. This set of parameters contain both physical and procedural effects on both data and voice communications networks.

Table 2.1: Parameter Descriptions

PARAMETER(S)	DEFINITION	NETWORK
Antenna Gain	A unitless number that relates the intensity of an antenna in a given direction to the intensity that would be produced by a hypothetical ideal antenna that radiates equally in all directions and has no losses.	DATA
Link Refresh - Distance	Measured in meters, refresh rate for network adaptations in terms of distance.	DATA
Link Refresh - Time	Measured in seconds, refresh rate for network adaptations in terms of time.	DATA
Transmission Control Protocol (TCP) Wait Time (Retransmit Interval)	Measured in seconds, wait time before trying to resend failed message.	DATA and VOICE
Max Transmission Control Protocol (TCP) Retries	Maximum number of attempts to send a failed message.	DATA and VOICE
Bits in Data Unit	Allowable size of message packet; Max Transmission Unit (MTU) is made up of the packet size and the size of the overhead.	DATA and VOICE
Bits of Overhead	Overhead in message packet.	DATA and VOICE
Available Bandwidth	A rate of data transfer, throughput or bit rate, measured in bits per second (bps).	DATA
Propagation Mode + Probability of Communications Success (PCOM)	Propagation is the behavior of radio waves when they are transmitted from one point on the Earth to another, or into various parts of the atmosphere while being affected by reflection, refraction, diffraction, absorption, polarization, and scattering. However, in S4 Propagation Mode is a *flag* between when this effect is "perfect" and when it has impact represented by PCOM.	DATA

CHAPTER 3: SCENARIO DESIGN

In order to evaluate a simulation, it is necessary to choose a scenario or scenarios that both represents the purpose of the simulation's intended use and emphasizes the parameters of interest in the thesis. This chapter describes the terrain, units in play, communication networks, and decision templates that make up the scenarios used for all experimentation.

3.1 SCENARIO CHOICE

The primary focus of this thesis is the interaction of communication environment parameters within the model. With that in mind, NMSU-PSL conceived and implemented a simple scenario containing a minimal number of agents and using well-traversed terrain from the National Training Center at Fort Irwin, California. To stress the communications networks, agents are required to almost continuously report their locations, in addition to observation reports and issued orders. Furthermore, to reduce variability, the DMP processes are kept simple and all ballistic engagements are removed in order to control the number of outcomes and ensure that the number of communication failures cannot be attributed to agent losses. As a result, the output of the simulation contains extensive data regarding communications with little confounding from other processes or events.

3.2 SITUATION

3.2.1 Agents

Agents within the scenario are split into two sides: BLUE and RED. BLUE is the predominant force in the battle space and the most complex. Shown in Table 3.1, BLUE contains 16 agents comprised of three commanders, two different types of platforms (High Mobile Multipurpose Wheeled Vehicles [HMMWVs] and an Unmanned Aerial Vehicle [UAV]), and dismounts. BLUE also contains five different DMPs: Battalion, Company, Recon, Platoon, and Platform.

Shown in Table 3.2, RED is much smaller, comprised of only five agents containing one commander and one type of platform (a generic truck). With fewer agents, RED also has fewer DMPs: Company and Platform.

Table 3.1: BLUE Force Agents

UNIT	AGENT	DESCRIPTION
Battalion Commander	BN CDR	Battalion DMP; provides C2 to entire BLUE force.
Maneuver Company Commander	CO CDR	Company DMP; provides C2 to the Maneuver Platoon; has its own HMMWV.
Reconnaissance Company Commander	Recon CDR	Recon DMP; provides C2 to the UAV and both Scout Teams.
Maneuver Platoon (PLT) Leader	HMMWV-4	Platoon DMP; relays information to and from the Maneuver PLT; has its own HMMWV.
Maneuver PLT	HMMWV-1 HMMWV-2 HMMWV-3	Platform DMP; each agent has its own HMMWVs.
Scout Team 1	Scout-1/1 Scout-1/2 Scout-1/3 Scout-1/4	Dismounts providing static overwatch in search of enemy activity; Scouts 1/4 and 2/4 serve as their respective "team leaders," relaying information to and from their teams.
Scout Team 2	Scout-2/1 Scout-2/2 Scout-2/3 Scout-2/4	
Unmanned Aerial Vehicle	UAV	Modeled as a "Shadow," provides aerial overwatch in search of enemy activity.

Table 3.2: RED Force Agents

UNIT	AGENT	DESCRIPTION
Insurgent Company Leader	Insurgent LDR	Company DMP; provides C2 to entire RED force; has its own Truck.
Insurgent Team 1	Truck-1 Truck-2	Platform DMP; each agent has its own Truck; Trucks 1 and 3 serve as their respective "team leaders," relaying information to and from their teams.
Insurgent Team 2	Truck-3 Truck-4	

3.2.2 Communications Networks

The scenario contains nine communications networks that are shown in Figure 3.1. Since agents can only communicate with other agents that are on the same network, each network is also a representation of the C2 structure for both forces. Each network is capable of handling both data and voice transmissions; however, each type of transmission requires different equipment in order to send and receive.

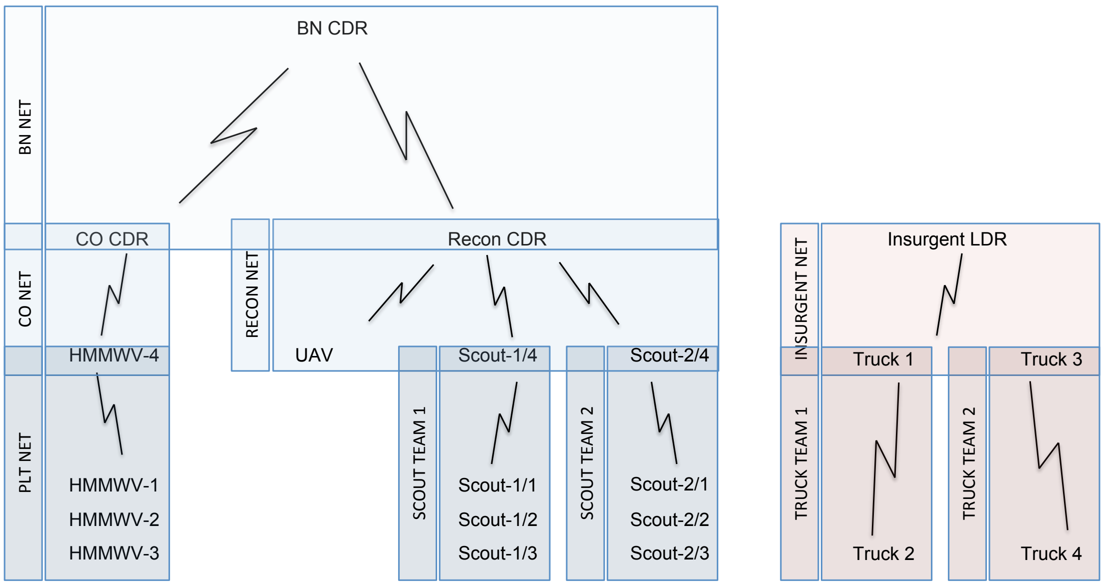


Figure 3.1: Graphical depiction of communications network in this thesis. The BLUE force network is depicted on the left and the RED force network is depicted on the right.

3.2.3 Concept of the Operation

Essentially, there is one scenario with three variants, each with the same concept. BLUE is operating on the eastern side of a mountain. Remaining with his Reconnaissance (Recon) CDR at the Assembly Area (AA), the BN CDR orders his CO CDR and Maneuver PLT to move toward a specified Objective (OBJ). To support their movement, BLUE also has its two Recon Teams providing overwatch along the two passes that go through the mountain range, with Recon Team 1 in the northern pass and Recon Team 2 in the southern pass. Additionally, departing from the AA, the Recon CDR directs his UAV to fly directly to a point northwest of the mountain range and then begin a counterclockwise overwatch of the entire range. To guide the reconnaissance effort, BLUE has designated an Area of Interest (AOI), made up of the two passes and the eastern side of range, in which BLUE does not want RED to enter.

Meanwhile, on the western side of the mountain, the RED Insurgent LDR orders himself and his two teams to move toward their own OBJ on the eastern side of the mountain range through the two passes. Moving through the northern pass is the Insurgent LDR and Insurgent Team 1 and moving through the southern pass is Insurgent Team 2.

Shown in Figure 3.2, the three variants change in the location of the AA and the BLUE OBJ. As a result, the direction the CO CDR, Maneuver PLT, and UAV move also change between variants. Everything else remains the same. There is one variant with the AA in the east and the OBJ in the north, one with the AA in the south and the OBJ in the north, and, lastly, one with the AA in north and the OBJ in the south. For this thesis, each variant is named and referred to based on the location of the AA: EAST, SOUTH, and NORTH.

3.3 SIMULATION EXECUTION

3.3.1 Reports and Messages

Within the scenario, there are four different types of communication. Listed in Table 3.3, the types differ from each other in function, means of transmission, and occurrence. Due to the high frequency of "spot" reports once BLUE forces locate and begin reporting on RED forces, there are more voice transmissions within this scenario than data. It is important to remember that only four of the selected parameters affect voice communications.

Table 3.3: Reports and Messages

NAME	TYPE	FREQ	DESCRIPTION
Situational Awareness (SA) Local Reports	Data	Every 62 seconds	Each agent reports what it knows about the other agents on the battlefield.
Situational Reports (SITREP)	Data	Every 106 seconds	Each agent reports location, current mission, movement, and equipment status.
Spot Reports	Voice	Dependent	Sent only when the sender is in visual contact with the enemy. Spot reports send information regarding what opposing agent they see, where it is located, and what it is doing.
MIP	Voice	Dependent	Sent only when the decision maker issues a change of mission.

3.3.2 Decision Templates

Table 3.4 specifies the decision templates employed in this scenario. Part of the simplified scenario was a reduced number of decisions to be made by each DMP. Again, the interest of this thesis is the communication process that drives the simulation's DMPs. Therefore, having

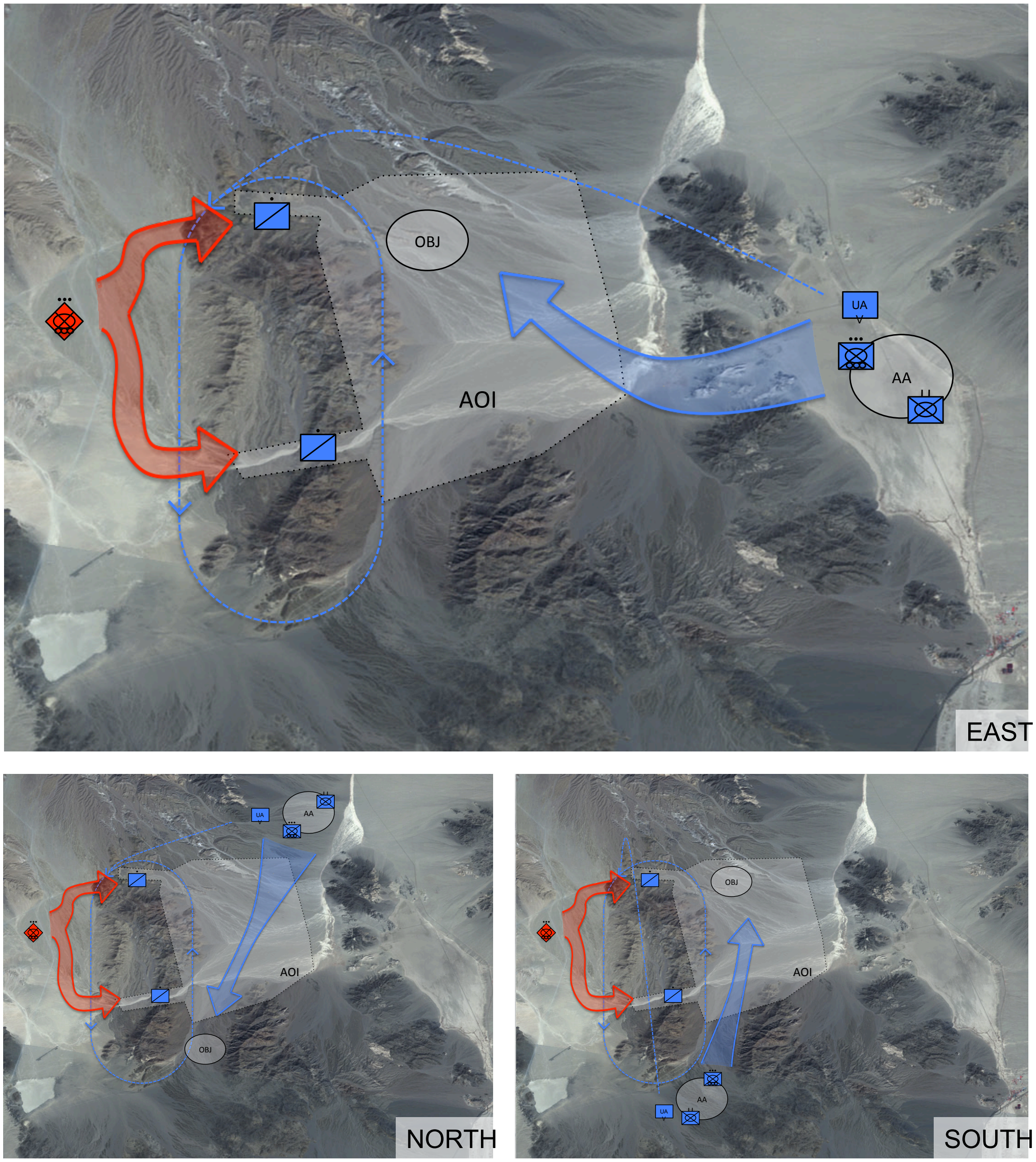


Figure 3.2: Graphical depiction of the scenario variants. Clockwise from top: EAST, SOUTH, NORTH.

complicated decision templates provides no additive value when all that needs to be seen is when rather than how a particular decision was made.

In general, BLUE's template is simply to send the CO CDR and his Maneuver PLT from the AA to the OBJ only in the absence of RED. As soon as RED is observed within BLUE's AOI, the CO CDR and Maneuver PLT are ordered to return to the AA or to the nearest designated rally point. Rally points are the AAs from the other scenario variants.

RED's template is even simpler: Cross the mountain through two passes, in two teams, toward a specified OBJ on the other side.

Table 3.4: Decision Templates

UNIT	EVENT	DECISION
BLUE	No enemy present (spotted) within the AOI	CO CDR and Maneuver PLT move from AA to OBJ.
	Enemy present (spotted) in AOI	CO CDR and Maneuver PLT change course and return to AA or nearest rally point.
	Enemy within engagement range	Do not engage enemy.
RED	Immediate	All RED forces move to OBJ; send Team 1 through northern pass, Team 2 through southern pass.

3.4 SUMMARY

This scenario was selected by NMSU-PSL to emphasize analysis of communication parameters. S4 is a large model with a vast set of parameters, ranging from the simple, such as terrain and movement, to the more complex ballistic and armor effects. A scenario that incorporates all the capabilities of S4 would be a great showcase for the simulation, but would seriously confound the parameters of interest in this thesis.

CHAPTER 4: EXPERIMENTAL DESIGN

The purpose of this thesis is to gain an understanding of the interactions and effects of our chosen parameters within S4. To accomplish this, this thesis follows the iterative data farming method described by Horne and Meyer (2004). Data farming allows for the exploration of an entire landscape of outcomes without having to explore every possibility. Data farming—given a model or scenario—is typically a four-step process:

1. Define parameters of interests.

2. Create a DOE to efficiently analyze the parameters at various values.

3. Conduct multiple runs of the simulation in accordance with the DOE.

4. Perform data mining on the results to obtain insight into relationships in the model.

Ideally, this process is iterated, with the finding from one set of experiments guiding the design of a subsequent set of experiments. This chapter explains that iteration was not possible for this study. Furthermore, having already defined the parameters, this chapter also describes the ranges of and the rules for our parameters, the methodology used to create the DOE, and the process in which the experimental runs were conducted.

4.1 PARAMETER RANGES AND RULES

For the purpose of this thesis, NMSU-PSL chose to analyze 10 communications parameters, comprised of continuous, discrete, and categorical values, believed to be the most important to the S4 communications environment. Not knowing the exact impact of changing these parameters, NMSU-PSL typically held them at constant values when using the model. In reality, the values of these parameters might increase or decrease (perhaps only slightly) from those typical values chosen by NMSU-PSL, so such an assumption cannot be guaranteed. Additionally, we can gain more insight on each parameter's impact by "stressing" it; that is, looking at ranges that include not just the typical cases, but the extreme ones as well.

The process used for choosing these "extreme" ranges (see Table 4.1) was more discretionary than scientific. For parameters where NMSU-PSL's typical or base value was somewhat of a median selection (*TCP Wait Time*, *Max TCP Retries*, and both *Link Refresh* parameters), the chosen ranges of study were 1 for the low end and a value twice as large as the base for the high end. For parameters where the base value was already at the high end (*Antenna Gain*, *Bits in Data Unit*, *Bits of Overhead*, *Available Bandwidth*, and *Propagation Mode + PCOM*), only a low end was chosen. Although the low ends chosen for *Bits in Data Unit*, *Bits of Overhead*, *Available Bandwidth*, and *Propagation Mode + PCOM* may not seem "extreme" in that they do not go to 1 or 0, they are considered "extreme" within S4 in that they place great stress on the modeled communications network. The low end for *Antenna Gain* also does not go to 1, but for a different reason. In S4, *Antenna Gain* contains three different values for three different groups of angles: $0°$ to $40°$, $41°$ to $80°$, and $\geq 81°$. In our use of the model, the range value for *Antenna Gain* determines the input value for angles $0°$ to $40°$. For angles $41°$ to $80°$, $\frac{1}{2}$ of the range value is used, and the value of 1 is used for angles $\geq 81°$. Following this rule, using the base value of 10 as the input value produces the three values of 10, 5, and 1 for $0°$ to $40°$, $41°$ to $80°$, and $\geq 81°$, respectively. Since it would not make sense to have a low angle range of 1, $\frac{1}{2}$, and 1, the low end was chosen to be 2, resulting in a low angle range of 2, 1, and 1.

Using these ranges, some initial trials were run that unearthed an unexpected relationship. In order to vary *Available Bandwidth*, another parameter, *Duration Mode* (time needed to transmit message), must be toggled between two values. This requirement both increased the number of parameters of interest to 11 and changed *Available Bandwidth* from a continuous to a categorical parameter, similar to *Propagation Mode + PCOM*. In addition to adding *Duration Mode*, there exist additional rules that govern the use of those two and other parameters. Some of the rules are pretty straightforward, but others are not. They are as follows:

1. To adjust *Bandwidth*, *Duration Mode* must be set to "2" (*Bandwidth* is "perfect" when *Duration Mode* is set to "0," meaning the message is instantaneously delivered).
2. To adjust *PCOM*, *Propagation Mode* must be set to "PCOM."
3. If *Propagation Mode* is set to "PCOM," then *Duration Mode* cannot be set to "0."
4. If *Propagation Mode* is set to "PERFECT," then *PCOM* cannot be set to any value other than 1.0. (However, when *Propagation Mode* is set to "PERFECT" the probability of success is 0.99).
5. The *Bits in Data Unit* value, plus the *Bits of Overhead* value, cannot exceed 4,608.

Table 4.1: Parameter Types and Ranges

PARAMETER(S)	TYPE	BASE	RANGE
Antenna Gain	Continuous	10	2 : 10
Link Refresh - Distance (Meters)	Continuous	20	1 : 40
Link Refresh - Time (Seconds)	Continuous	20	1 : 40
TCP Retransmit Interval	Continuous	1.0	0.1 : 3.0
Max TCP Retries	Discrete	3	1 : 6
Bits in Data Unit	Discrete	4,448	2,704 : 4,448
Bits of Overhead	Discrete	160	160 : 624
Duration Mode + Available Bandwidth	Categorical	0 + 1.0	0 + 1.0 2 + 0.9 2 + 0.8 2 + 0.7 2 + 0.6
Propagation Mode + PCOM	Categorical	PERFECT + 1.0	PERFECT + 1.0 PCOM + 0.9 PCOM + 0.8 PCOM + 0.7 PCOM + 0.6 PCOM + 0.5

The first four rules were handled by grouping the values for *Duration Mode*, *Available Bandwidth*, *Propagation Mode*, and *PCOM* into one categorical parameter comprised of all 36 possible combinations allowed by the rules. For the last rule, a constraint was placed on the two parameters, preventing their addition from exceeding the max allowed value. It should be noted that *Bits in Data Unit* and *Bits of Overhead* normally have a dependent relationship in that the addition of the two typically adds up to 4,608 in order to maximize the total number of bits available for each message. However, enabling "under-usage" went toward the intent to stress these two parameters.

4.2 DESIGN OF EXPERIMENT (DOE) SELECTION

To efficiently analyze S4 in the context of our parameters, an efficient design is needed. To simply look at every possible combination between the parameters, this experiment would take quite some time, to say the least. For example, to analyze nine parameters (ignoring the parameter rules for the purpose of this exercise), using a DOE made up of the combinations produced by a traditional two-level factorial design that looks at just the "high" and "low" values for each, 2^9 or 512 design points would need to be run (Sanchez, 2008). Such an experiment would provide great information at what is occurring at the corners of the model, but despite the large number of design points, would provide little insight as to what is occurring in the middle of that design space. In other words, there would be great insight into what is happening to the model when *Bits in Data Unit* is at 2,704 and 4,448, but not when that parameter is at, say, 2,800, 3,202, or at any other value in between. Unless the response to each parameter and their interactions are assumed to be linear or of some other functional form, something needs to be known about what is happening throughout the design space. Following the factorial or combinatorial approach to sample in the middle would dramatically add to the number of experiments. For example, a five-level, full-factorial design would produce 5^9 or nearly 2 million design points (Sanchez, 2008). What we want instead is a design that both fills the design space and requires fewer experiments to run.

4.2.1 Nearly Orthogonal Latin Hypercubes (NOLH)

The development of NOLH by Cioppa and Lucas (2007) expanded on the earlier work of Orthogonal Latin Hypercubes (OLH) by Ye (1998) and provided experimenters with readily available efficient, space-filling, and nearly orthogonal designs for continuous parameters. These designs are efficient in that they explore many factors in a modest number of experiments. The fact that they are space-filling means that the designs sample throughout the feasible region. The nearly-orthogonal nature reduces the maximum absolute pairwise correlation (ρ_{map}) between the columns of the design matrix, which allows for analysis that is not plagued by adverse effects, such as confounding, due to multicollinearity (Montgomery, Peck, & Vining, 2006). Moreover, these designs provide analytic flexibility by allowing experimenters to fit a broad range of potential meta models and visually examine many relationships (Kleijnen, Sanchez, Lucas, & Cioppa, 2005).

With only nine parameters for this thesis, an NOLH is available that requires only 33 design points (see http://harvest.nps.edu). With this design, an apparently great space-filling design is created. However, the max ρ between our parameters is 8.9%. A design is perfectly orthogonal when $\rho_{map} = 0$ and is considered nearly-orthogonal when $\rho_{map} < 5\%$. This high ρ_{map} highlights the main shortfall of the NOLH. While the design is great at handling continuous values, as it was designed, it is limited when it comes to discrete and categorical values. In order to handle discrete values, the NOLH software rounds its calculations, thus creating error and increasing

ρ_{map}. In order to handle categorical values, discrete numbers are substituted for each category. This adds error from simply treating the values as discrete, but also because it forces a mathematical distinction between categories that more than likely does not exist. For instance, to handle a category that contains the colors RED, BLUE, and GREEN, the NOLH would treat them as a category that ranged from $1:3$. Substituting 1 for RED, 2 for BLUE, and 3 for GREEN, the NOLH would produce a design that rounded orthogonal values for 1, 2, and 3. But do those numerical values necessarily map to our three colors? Is BLUE really one value and GREEN two values away from RED?

Since the objective is to obtain a ρ_{map} as close to 0 as possible, the NOLH design allows for other operations to mitigate the correlation created by discrete and categorical induced errors. This is done by increasing the number of design points by either "stacking and rotating" smaller designs or applying a design intended for a greater number of parameters on a smaller set. Applying the latter operation, a 256-design point NOLH, which is capable of handling up to 29 parameters, produces a reduced ρ_{map} of 1.9%. However, this operation is done at the expense of an additional 223 design points—a much greater computing cost.

4.2.2 Nearly Orthogonal Nearly Balanced Mixed Design (NONBMD)

This issue, created by discrete and categorical parameters, has been addressed through the work of Vieira and Sanchez (2010). They developed an efficient NONBMD using a mixed-integer program that, in addition to utilizing orthogonality and minimizing ρ_{map}, takes into account the imbalance of an experiment by addressing mixed factor types of varying ranges. Since this approach varies based on the parameter and ranges under examination, there are currently no catalogued designs that an experimenter can simply grab and use. Fortunately, Vieira was able to create a custom NONBMD for this thesis that incorporates all of the parameters, ranges, and rules. The result of his custom design is a 144-design point DOE that has both more space-filling than the NOLH designs and a lesser ρ_{map} of 1.4%. Although this is only slightly better than our 256-design point NOLH, it is a design nearly half the size! Due to the size of the *Duration Mode + Available Bandwidth + Propagation Mode + PCOM* categorical parameter and the large range of *Bits in Data Unit*, the number of design points could not be reduced further without increasing ρ_{map}. (To see the comparisons between the design correlations and design spaces, see Figures 4.1 and 4.2. To see the full NONBMD design matrix, see Figure 4.3. The dots in Figures 4.2 and 4.3 represent the input values for pairs of parameters.)

NONBMD Correlation Matrix (144 DPs): Max Pair-wise Correlation = 0.0138

	Grouped Category	Antena Gain	Max TCP Retries	TCP Retrans Interval	Link Refresh Meters	Link Refresh Seconds	Bits Data Unit	Bits Overhead
Grouped Category	1.0000	−0.0000	−0.0000	−0.0000	0.0016	−0.0022	−0.0128	−0.0138
Antena Gain	−0.0000	1.0000	−0.0000	−0.0002	0.0021	−0.0018	0.0000	0.0000
Max TCP Retries	−0.0000	−0.0000	1.0000	−0.0004	0.0020	0.0000	−0.0000	0.0000
TCP Retrans Interval	−0.0000	−0.0002	−0.0004	1.0000	0.0006	−0.0017	−0.0002	0.0003
Link Refresh Meters	0.0016	0.0021	0.0020	0.0006	1.0000	0.0010	0.0020	0.0019
Link Refresh Seconds	−0.0022	−0.0018	0.0000	−0.0017	0.0010	1.0000	0.0004	0.0006
Bits Data Unit	−0.0128	0.0000	−0.0000	−0.0002	0.0020	0.0004	1.0000	−0.0090
Bits Overhead	−0.0138	−0.0000	0.0000	0.0003	0.0019	0.0006	−0.0090	1.0000

NOLH Correlation Matrix (33 DPs): Max Pair-wise Correlation = 0.0889

	Grouped Category	Antena Gain	Max TCP Retries	TCP Retrans Interval	Link Refresh Meters	Link Refresh Seconds	Bits Data Unit	Bits Overhead
Grouped Category	1.0000	−0.0300	0.0338	−0.0242	−0.0243	−0.0154	0.0280	0.0018
Antena Gain	−0.0300	1.0000	0.0889	0.0294	0.0158	0.0061	0.0095	−0.0309
Max TCP Retries	0.0338	0.0889	1.0000	−0.0357	−0.0698	−0.0798	0.0245	0.0122
TCP Retrans Interval	−0.0242	0.0294	−0.0357	1.0000	−0.0005	−0.0024	−0.0052	−0.0134
Link Refresh Meters	−0.0243	0.0158	−0.0698	−0.0005	1.0000	0.0027	−0.0005	−0.0164
Link Refresh Seconds	−0.0154	0.0061	−0.0798	−0.0024	0.0027	1.0000	0.0049	0.0246
Bits Data Unit	0.0280	0.0095	0.0245	−0.0052	−0.0005	0.0049	1.0000	−0.0020
Bits Overhead	0.0018	−0.0309	0.0122	−0.0134	−0.0164	0.0246	−0.0020	1.0000

NOLH Correlation Matrix (257 DPs): Max Pair-wise Correlation = 0.0194

	Grouped Category	Antena Gain	Max TCP Retries	TCP Retrans Interval	Link Refresh Meters	Link Refresh Seconds	Bits Data Unit	Bits Overhead
Grouped Category	1.0000	−0.0010	0.0082	−0.0035	−0.0022	0.0013	0.0017	0.0037
Antena Gain	−0.0010	1.0000	−0.0105	0.0107	0.0013	0.0099	0.0010	0.0061
Max TCP Retries	0.0082	−0.0105	1.0000	−0.0079	0.0094	−0.0089	0.0102	−0.0194
TCP Retrans Interval	−0.0035	0.0107	−0.0079	1.0000	−0.0016	0.0018	0.0032	0.0010
Link Refresh Meters	−0.0022	0.0013	0.0094	−0.0016	1.0000	0.0000	−0.0028	−0.0061
Link Refresh Seconds	0.0013	0.0099	−0.0089	0.0018	0.0000	1.0000	0.0008	0.0027
Bits Data Unit	0.0017	0.0010	0.0102	0.0032	−0.0028	0.0008	1.0000	−0.0001
Bits Overhead	0.0037	0.0061	−0.0194	0.0010	−0.0061	0.0027	−0.0001	1.0000

Figure 4.1: The NONBMD (top) has a lower max pairwise comparison (ρ) than both the 33-design point NOLH and the 256-design point NOLH. Max ρ for each design is circled in red.

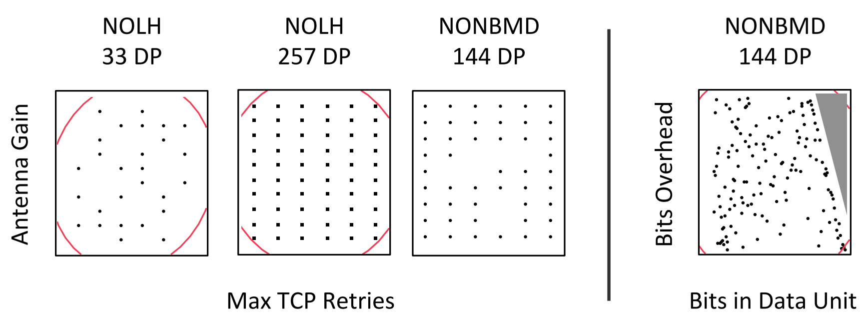


Figure 4.2: NONBMD (third plot from left) provides tremendous design space coverage over the NOLH. Note the corner lines in each of the design space blocks for the 33- and 256-design point NOLH designs depicting the reduced number of samples near the corners. The plot at the far right shows the impact of the "Bit" restriction on the Design Space. The shaded triangle is the area in which rule number 5 is violated and therefore, no sampling is done in that area.

4.3 REPLICATION

Random numbers are used within S4 to handle the various probabilities that are modeled, e.g., determining the probability of communication success manifested in the parameter *PCOM*. Consequently, runs with all of the same input parameter values, but with different initial random seeds, will produce different results. Replication of each design point then becomes a necessity

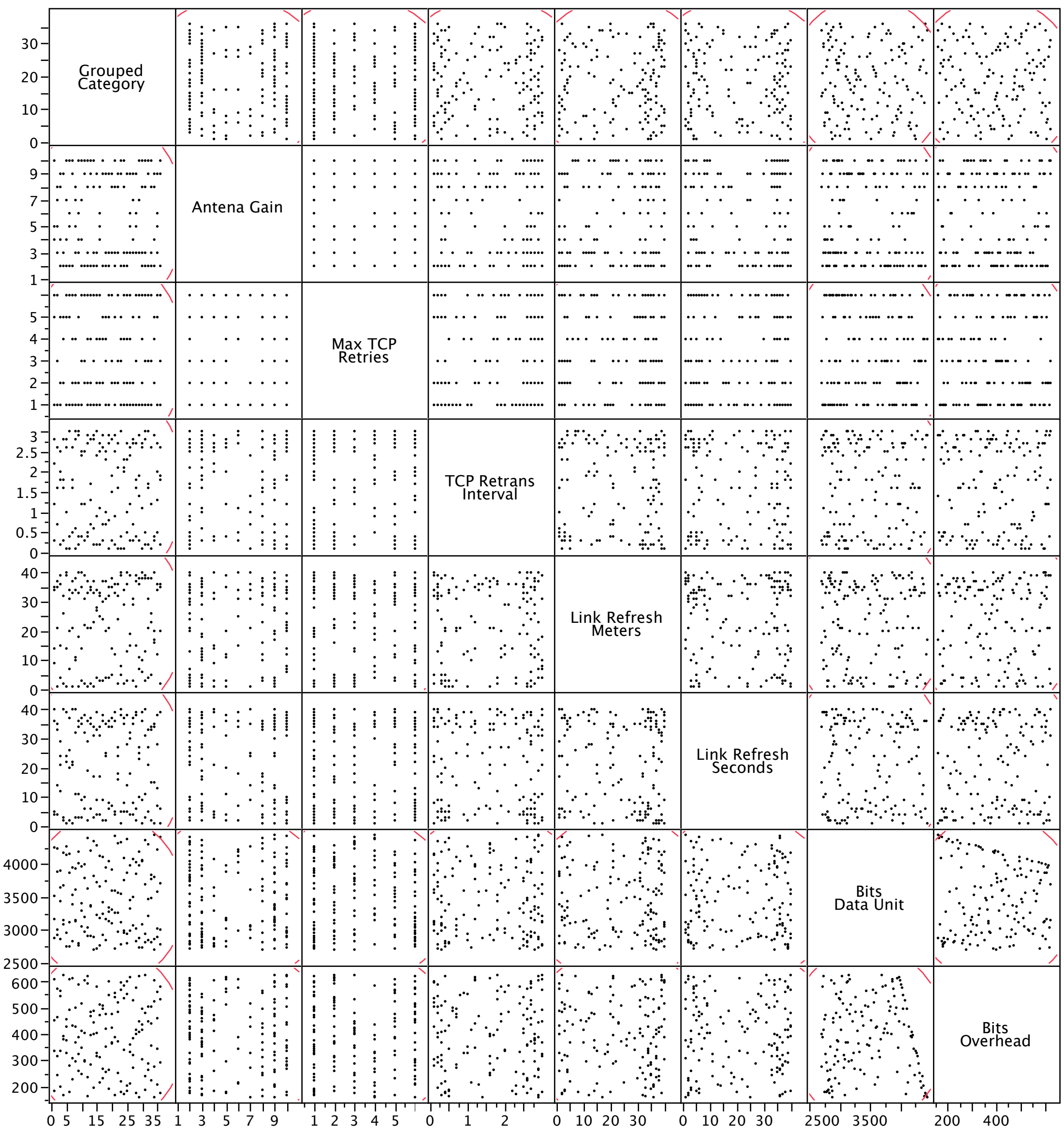


Figure 4.3: NONBMD provides tremendous design space coverage over the NOLH and easily incorporates all the parameter rules. See Appendix A for corresponding matrices for both the 33- and 256-design point NOLH.

in order to provide multiple results that will enable statistical inferences to be made. In order to determine how best to utilize replications, there must first be an understanding of how random numbers are used within S4.

As previously mentioned, S4 takes one initial random seed to generate one stream of random numbers from which multiple probability processes, such as communication, identification, and ballistic impacts, draw. To date, NMSU-PSL does not have a firm grasp on when and in what order each random number is used (although this is something they are currently working on). This creates a problem with design point replication: with multiple processes using the same stream of random numbers and not knowing when those numbers are being drawn, if the same initial random seed is used with different parameter values, the difference between the results cannot be guaranteed to only be a function of the changes in the parameters and independent random error. This is because that the runs cannot be guaranteed to be independent—especially if critical events take place early in the simulation. If the processes used separate streams, variance reduction techniques could be applied that would increase the statistical power and/or reduce replication requirements (Law & Kelton, 2000). To incorporate separate streams would require redesigning S4; therefore, to mitigate this issue, a different initial seed will be used for every run.

Next, the number of replications must be determined. To find this number, the variance created from only changing the initial seed within a single, representative design point must be calculated. To aid in this endeavor, NMSU-PSL conducted 10 "Seed Variance" runs using their base parameter values (see Table 4.1) for each run while changing the initial random seed from 1 to 10. Using total number of communications failures as the response, a Q-Q plot of the results, shown in Figure 4.4, produced an interesting picture. While the first eight runs appeared to follow nicely along the normal line, the last two runs lay way above it. Not knowing if this occurrence is a function of the simulation or merely a fluke, NMSU-PSL produced two more runs using seeds 11 and 13. These two additional runs resulted in total failures similar to runs 9 and 10, thereby pulling the normal line up and through all 12 runs. When plotting the total failures from all 12 runs for just one agent—the UAV—the results fit even better along the normal line. Although far from convincing, given our small sample and limited purpose for doing so, it seems reasonable to suggest the response is approximately normally distributed.

In order to calculate the number of replications, the following "power" formula was used: $n = \frac{2Z_{\alpha/2}\sigma}{w}$, where n is the number of replications per design point and w is the width of the confidence interval surrounding the results. As w decreases, the confidence of the results increases. Using this calculation, our assumption of normality and the variance produced from our 12 Seed Variance runs, using total communications failures as the response ($\hat{\sigma}^2$ = 11.9), the plot shown in Figure 4.5 is produced.

Without the power calculation, it is obvious that as n increases, w decreases. However, from the plot of the calculation, the point of diminishing returns is shown to exist when n is between 15 and 25. Such information is beneficial, as each replication requires additional computing resources. In the case of this thesis, this calculation is instead used to show what will be the resulting confidence level. The NONBMD design that greatly improved space-filling and reduced correlation from the 33-design point NOLH, did so at the expense of adding an additional 111 (total 144) design points. Therefore, without any replication, our experiment already consists of 432 runs (144 design points $*$ 3 scenarios). Having chosen to make the runs themselves, NMSU-PSL would be using only one computer—and not a cluster of computers—to run the entire experiment. With each run taking just over two minutes to complete, adding too many replications would create an excessive burden. After balancing the computing power and time needed to conduct the runs with the need to have replication and increased parameter ranges to stress the model, the number of replications was limited to five, for a final tally of 2,160 runs. Therefore, the 95% confidence w for this thesis will be approximately 6.

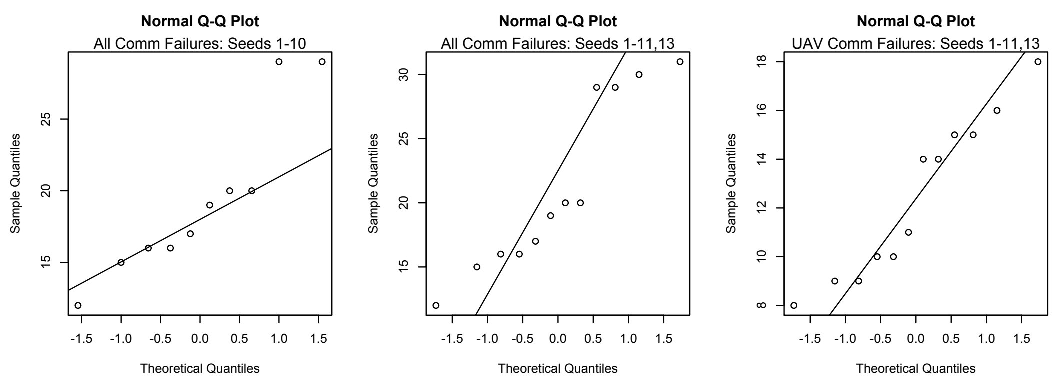


Figure 4.4: The initial Seed Variance runs produced two values of total communications failures that did not follow the normal line (left plot). After inputting results from two additional runs (middle plot), the normal line was pulled higher, but the results are not great. Taking samples from just one agent (right plot), however, does produce results that seem to follow a normal line.

4.4 RUN EXECUTION

The processes of creating the simulation execution files and conducting the experiment runs themselves were both simplified with the use of computer programs or scripts. S4 is designed so that a Graphical User Interface (GUI) can be used to select and input each parameter value within the model. Fortunately for this study, a programmer can also manipulate the parameters within an XML "execution" file. To conduct this experiment, 432 execution files had to be made with the parameter values for each communications device or agent changed in each file,

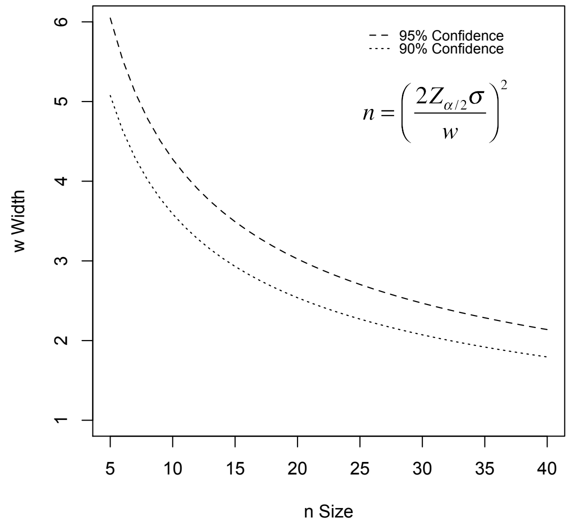


Figure 4.5: Power Calculations using total communications failures as the response. As n increases, w decreases. Future studies of S4 should try to incorporate more replications at each design point.

in accordance with the DOE. Accomplishing such a task manually would be quite tedious, time consuming, and prone to error. An NPS Research Associate, Steve Upton, developed the program *XStudy* that maps the parameter values of a DOE in an Excel spreadsheet to a simulation's execution file. Once all the locations of the parameters within the execution file are identified, the process of creating an entire study's execution files is reduced from weeks to hours. (For more information on *XStudy* see http://harvest.nps.edu/.)

Since S4 was not built to conduct numerous runs simultaneously and access to cluster computing was unavailable for this thesis, so to execute the runs NMSU-PSL produced a simple script that read into the simulation each execution file five times, incrementing the initial seed by one from one run to the next. Under this approach, each design point was replicated five times before the next design point was read in. A computer cluster of a few dozen processors could have then completed the experiment runs overnight and with more replications. Limited to the use of a sole computer after close of business and various times during the day, NMSU-PSL conducted all 2,160 runs in over 80 hours completed over the course of seven days. Due to the size of the output data (approximately 222 gigabytes [GBs]), transfer from NMSU-PSL to NPS was unsuccessful using either the Internet or cloud storage disks. In order to complete the transfer, the output was stripped of all files deemed unnecessary for this thesis as well as one file that was just too large. The remaining files (approximately 19 GBs) were then compressed and burned onto four data digital video discs (DVDs) and shipped via FedEx. This data compression process corrupted two of the run files: one from the EAST variant and one from the SOUTH; both from different design points. Although these run files were subsequently transferred as complete files, they arrived after the analysis had already been completed. As a result, instead of 2,160 runs, 2,158 were analyzed for this thesis.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 5:
DATA ANALYSIS

This chapter explains the fourth step to the data farming process: the reasoning behind the selection of the response variables, as well as the methods by which they were extracted from the simulation's output. Furthermore, it explains the process of how the output data was examined using data farming tools and statistical software in order to determine if any or all of the parameters in question, and their interactions, are providing unintended results or behaviors.

5.1 RESPONSE SELECTION AND COLLECTION

It has been repeatedly mentioned that S4 produces a large amount of output data. Collecting the data for each event—and in some cases, each time step—for various processes such as communication, decision making, and movement for this scenario creates a roughly 92 megabyte (MB) output file for one experimental run. Such voluminous output enables NMSU-PSL and ARL-SLAD to examine the results of a single run from many different angles. For the number of runs that this thesis is to analyze and in the time available, however, analysis must be limited to only a few select results or responses. One response is obvious: the immediate impact of the communications environment on communications success can be analyzed by the number of communications failures. Additionally, failures can be analyzed by both type and affected agent. In order to analyze the effect of those failures on the scenario's mission execution, a second response is needed. This second response must look at the time delay between when the RED Insurgent Teams are first reported as being in the AOI and, as a result of that occurrence, when the BLUE Maneuver PLT receives the MIP that changes its current mission. This second response will be referred to as the "Operational Delay."

5.1.1 Communications Failures

One of S4's output folders is "CommsDMPAudit." Within this folder is a file for each agent in the simulation that records every successful or attempted communication received and sent by that agent. When a message failure occurs, the word "REASON" appears, followed by the failure codes below. The first two codes refer to both voice and data failures and the second two refer only to voice failures:

- SEND_BY_EXPIRED
- NO_HEADER_ACK
- NO_DATA_ACK
- NET_CONTENTION

To extract this data, a script was written using Ruby coding language that opened up each agent file within the CommsDMPAudit folder and counted each instance of each failure. Additionally, to ensure that there were no other failure codes and that our search was done correctly, the word "REASON" was also counted to serve as a count of total failures. The output resulting from this code produced a file listing the number of failures by type and by agent. A quick check of the data showed that the sum of the failures by type equaled the sum of the total failures, thereby providing strong evidence that the search was correct.

5.1.2 Operational Delay

Within our scenarios, there is no flag set to identify when a RED agent is first spotted within the AOI. The report sent up from this event looks just like any other; it simply contains the perceived location of the RED agent. When the BN CDR receives this information, it compares the reported location with what it knows to be BLUE's AOI. Tracing down this communication chain presents many challenges and requires the use of multiple files. Simplifying it to the time when a RED agent enters into the AOI requires searching through the "Playback" file in order to know the exact location of each agent at each time step. Unfortunately, this file alone is 20 MB and was not available for this thesis. Calculating our Operational Delay, therefore, requires some creativity and a critical assumption.

If it is assumed that changes in communications do not affect the physical movement of the RED agents, it can be inferred that once each RED Insurgent Team begins movement, each of those teams will reach the AOI within the same amount of time. Of course, as explained in Chapter 4, the change of random seeds will prevent this time from being exact. The creativity will be in extracting not only when this movement begins, but also when the BLUE Maneuver PLT receives the MIP to change its mission. Returning to the CommsDMPAudit folder (for the sake of simplicity), after some trial and error, the following message key word sets are found that specify the events of interest:

- Insurgent Team begin movement: MOVE_AND_RETURN_FIRE
- Maneuver Platoon change of mission: CO-A-HQ to HMMWV-R4. Blue-Retreat

The first set of key words for the Insurgent Teams are extracted from Trucks 2 and 4. Therefore, the time step that this code is first received should be the time that both trucks in each team are aware of the movement order. The second set of key words for the Maneuver PLT are extracted from HMMWV-4, the PLT Leader. Receipt of this message means that the BN CDR has already made its decision to change the mission and issued the order to the Maneuver CO CDR. Since this message is extracted from HMMWV-4, the time step of its occurrence is only that the order has been issued, not when the platoon actually changes direction. In an attempt to ensure that these messages are only identified when they are successfully received, they must be preceded by the key phrase "ReceivedCommuniquePart." Unfortunately, if the message contained additional parts, this phrase does NOT guarantee that the entire message was received and acknowledged, only the part immediately following. To extract these time steps, a script was again written using Ruby coding language that opens up the specific agent files within the CommsDMPAudit folder, finds the occurrence of the key words, and returns the time step that it occurred for each Insurgent Team and the Maneuver PLT.

Since the change of mission should occur after the first Insurgent Team is spotted, each team's arrival to the AOI is calculated using the time step when each initiated movement, plus their corresponding average travel time. The smallest of those two calculations is then subtracted from the time step when the Maneuver PLT received the MIP in order to calculate the Operational Delay. To determine each team's average travel time and calculate an average Operational Delay time, the data extraction code was used on the Seed Variance runs. Then, manually going through the "Playback" files that were available for each of the 12 runs, the time step was recorded when an agent from each team passed a specified X Coordinate on or about the border of the AOI, the corresponding time step was recorded. Subtracting those time steps from the extracted "begin movement" time steps produced a mean travel time for Insurgent Team 1 of 21.15 minutes, with a 95% Confidence Interval of the mean of [20.58, 21.72], and 15.42 minutes, with a 95% Confidence Interval of the mean of [15.09, 15.75] for Insurgent Team 2. Using the extracted change of mission times and the Operational Delay formula produced a mean Operational Delay of 12.08 minutes, with a 95% Confidence Interval of the mean of [11.79, 12.37] for the Seed Variance runs.

5.2 MESSAGE "LISTENING"

As mentioned in Chapter 4, prior to building the DOE, some initial trials were conducted that varied the parameters within our stated ranges. These trials revealed an additional insight into the simulation that, although it had no impact on the DOE, is important to know when evaluating the results. Recall that Table 2.1 showed that five of the selected parameters affect only data communications and consequently have no impact on voice transmissions. As noted previously, the scenario in this thesis is heavier on voice communications since the majority of communications appear to be "spot" reports (an agent reporting the location and activity of enemy agents), which are voice. In an attempt to remove this imbalance by making Spot reports data instead, resulting runs revealed that data spot reports would go up the chain of command, but would never come back down. For example, if Scout-1/3 reported to his team leader Scout-1/4 that he spotted a RED agent, Scout-1/4 would relay that message up to the Recon CDR, but would not distribute that information to the rest of his team. Scouts-1/1 and 1/2 would have no idea that Scout-1/3 saw a RED agent. It turns out that on a voice network, the other agents on that network "listen" in on all communications that occur on that network. For a radio network, this seems like a realistic assumption for the simulation to make. However, on a digital device this "listening" ability would be dependent on the type of device and, therefore, the same assumption is not made for a data network. As a result of this realization, Spot reports remained as voice transmissions.

As mentioned previously, S4 outputs a "Perception" file for each agent that, among other things, tracks when an agent physically identifies another agent and when it is simply aware of another agent through its communications network. Due to listening, these files could not be used to help calculate the Operational Delay. When the Recon CDR sends his report of insurgent locations to the BN CDR, the Maneuver CO CDR hears the message traffic and relays that information to its Maneuver PLT. Therefore, the Maneuver PLT typically is aware of the presence of insurgents before the BN CDR issues the MIP.

5.3 COMMUNICATION FAILURES

Before analyzing the data, there are already some expected outcomes. From our knowledge of S4 and the DOE it seems obvious that we would expect to see an increase in failures as the communications environment becomes more restrictive. Additionally, based on results from the Seed Variance runs, there should be different distributions of failures among failure types as well as agents. Furthermore, based on the assumption used to develop our Operational Delay response, it might be expected that there is no significance in failures between scenario variants. However, since the changing location of the BLUE Assembly Area creates different distances between both the BN CDR and Recon CDR from the orbiting UAV and scout teams, a difference in failures could exist.

5.3.1 Failure Distributions

All of these expectations can be verified or rejected with some simple analysis. Producing a histogram of "Total Failures" regardless of type or agent (Figure 5.1) produces a skewed result that is clearly not normal. The mean number of Total Failures is approximately 84.9, with a 95% Confidence Interval of the mean of [83.34, 86.46] and a median of 77. The minimum number of failures in one design point is 4 and the maximum number of failures observed in another is 250. The bulk of the failures occurred between 60 (1st Quartile) and 105 (3rd Quartile). Breaking the failures down by type reveals that approximately 77% of the failures observed were of type SEND_ BY_ EXPIRED, whereas only 17% were NO_ HEADER_ ACK, 5% NO_ DATA_ ACK, and 1% NET_CONTENTION. These percentages makes some sense since SEND_ BY_ EXPIRED and NO_ HEADER_ ACK capture both voice and data failures, whereas the other two failures types only capture voice. However, the later three failure types produced histograms that are extremely skewed left, looking almost exponential in distribution, whereas SEND_ BY_ EXPIRED produces a distribution similar to Total Failures (Figure 5.2). The skewness of these failures could be a result of the fact that the parameter ranges are not all centered around their base values. Its possible that different ranges may produce results that are more normal.

A quick look at boxplots that separate failures by scenario variant gives the impression that there is not much difference between variants (Figure 5.3). However, a more detailed analysis, using a One-Way Analysis of Variance (ANOVA), reveals that although there is, in fact, no significant difference between variants for either NO_ HEADER_ ACK or NO_ DATA_ ACK, there is a significant difference in both Total Failures and SEND_ BY_ EXPIRED failures, and a slight difference in NET_ CONTENTION (Figure 5.4). Most notable is the fact that EAST has the lowest mean and median failures. Measuring from the scout teams and the center of the UAV's orbit, the AA in the EAST variant is farther away (approximately 18.7 km, 17 km, and 17.7 km, respectively) than both the NORTH (approximately 9.9 km, 13.6 km, and 11 km, respectively) and SOUTH variants (approximately 15.2 km, 7.4 km, and 11.6 km, respectively). If our earlier assumption that distance between the scout teams and UAV from the AA is a potential cause for any difference in failures between variants is true, then the EAST variant should have the highest mean. Clearly this is not the case. This could possibly be answered by looking at failures by agent.

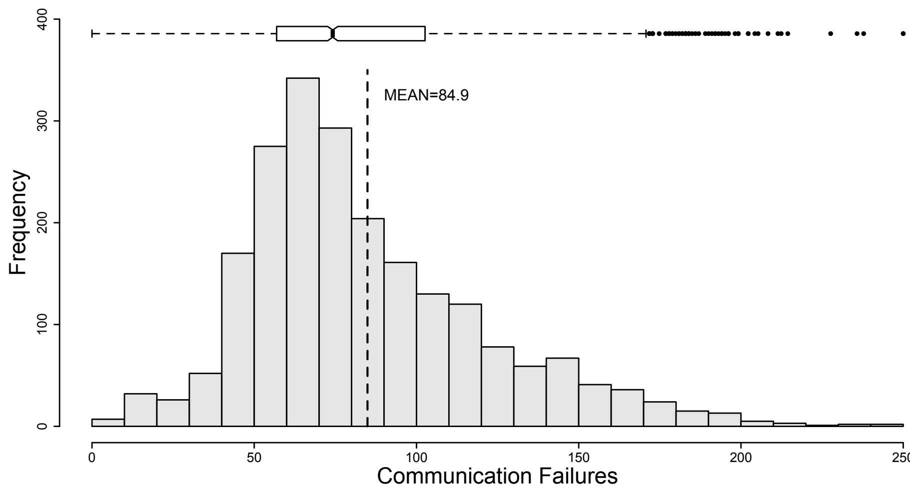


Figure 5.1: Histogram of Total Failures regardless of type or agent with corresponding boxplot. The distribution is clearly left skewed and not normal.

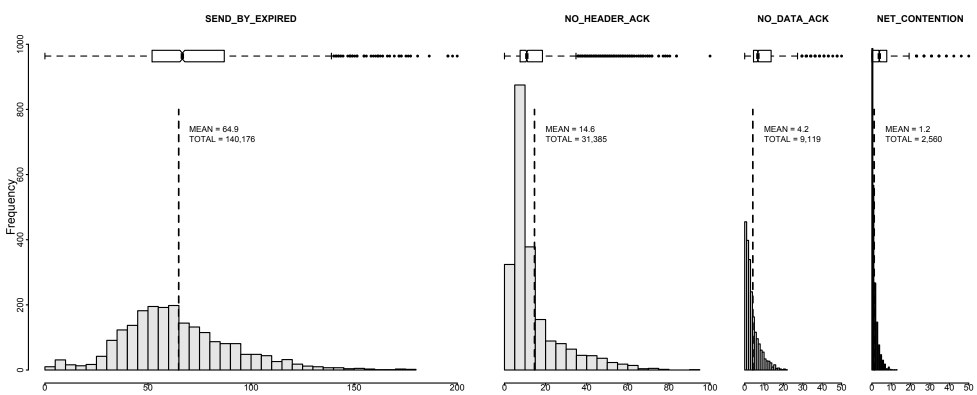


Figure 5.2: Histogram of Total Failures by failure type with corresponding boxplots. Clearly, SEND_ BY_ EXPIRED is the predominant failure type and has a distribution similar to Total Failures seen in Figure 5.1.

The Seed Variance runs produced failure numbers that predominantly came from the UAV and occasionally from various scouts. All other agents in those 13 runs had none. Figure 5.5 shows that the UAV is again the primary source of communications failures in our experiment; however, the HMMWV-4 PLT Leader claims the next highest number of failures instead of any of the scouts. An interesting result is the difference in failures between the two scout teams. The number of failures between Scout 1/4 (Team Leader) and Scout 1/1 are very close to the failures of Scout 2/4 (Team Leader) and Scout 2/1, respectively. However, the same cannot be said between the other members of the scout teams. Also worth noting is that Insurgent Truck 2 had zero failures in 2,159 runs. Other than these two interesting results, considering the scenario and each agent's DMP, the remaining numbers make sense. Agents with extremely low failures are either primarily receivers of information, sending communications primarily to acknowledge message receipt, or are issuers of occasional orders. Additionally, many of these agents' messages are sent to agents with whom they are in close proximity.

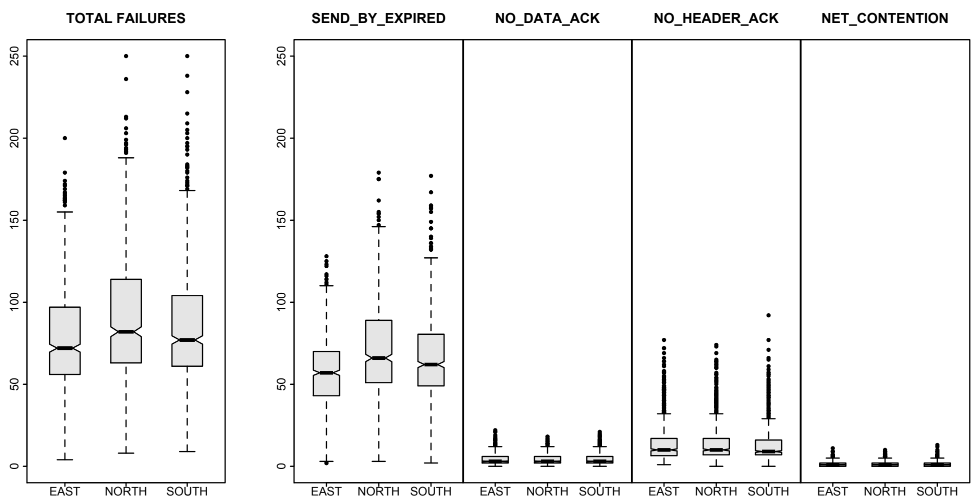


Figure 5.3: Boxplots by scenario variant of Total Failures and failures by type. Although similar in appearance, there is a significant difference between scenarios in Total Failures, SEND_ BY_ EXPIRED, and NET_ CONTENTION.

The difference between scout teams is suspected to be a result of the path traveled by Insurgent Team 1. If the insurgents did not travel exactly through the path on which two scouts from Scout Team 1 were on either side, the members of that scout team would not have as many observations of the insurgents to report. This would result in fewer sent messages and, consequently, fewer failures. However, this hypothesis was somewhat disproven by NMSU-PSL. Looking back at the "playback" output files, although there was a slight deviation from the path, NMSU-PSL did not believe it to be significant enough to create a difference in reports. Unfortunately, while it makes sense that Insurgent Truck 2 would have fewer failures than either Trucks 1 or 3 due to its position in the network, since its parameter values were varied in the same manner as every other agent, no clear explanation can be provided as to why it had 0 failures.

TOTAL FAILURES

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Scenario	2	66954.6	33477.3	25.1317	<.0001*
Error	2155	2870614.7	1332.1		
C. Total	2157	2937569.3			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
EAST	719	77.6398	1.3611	74.971	80.309
NORTH	720	91.1708	1.3602	88.503	93.838
SOUTH	719	85.9166	1.3611	83.247	88.586

Std Error uses a pooled estimate of error variance

SEND_BY_EXPIRED

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Scenario	2	65386.0	32693.0	47.0829	<.0001*
Error	2155	1496369.9	694.4		
C. Total	2157	1561755.9			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
EAST	719	57.7024	0.98272	55.775	59.630
NORTH	720	71.0306	0.98204	69.105	72.956
SOUTH	719	66.1280	0.98272	64.201	68.055

Std Error uses a pooled estimate of error variance

NET_CONTENTION

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Scenario	2	10.5395	5.26975	1.9253	0.1461
Error	2155	5898.5745	2.73716		
C. Total	2157	5909.1140			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
EAST	719	1.13908	0.06170	1.0181	1.2601
NORTH	720	1.13472	0.06166	1.0138	1.2556
SOUTH	719	1.28512	0.06170	1.1641	1.4061

Std Error uses a pooled estimate of error variance

NO_DATA_ACK

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Scenario	2	8.229	4.1144	0.3102	0.7334
Error	2155	28586.869	13.2654		
C. Total	2157	28595.098			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
EAST	719	4.14882	0.13583	3.8824	4.4152
NORTH	720	4.30000	0.13574	4.0338	4.5662
SOUTH	719	4.22809	0.13583	3.9617	4.4945

Std Error uses a pooled estimate of error variance

NO_HEADER_ACK

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Scenario	2	78.68	39.338	0.2336	0.7917
Error	2155	362968.73	168.431		
C. Total	2157	363047.41			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
EAST	719	14.6495	0.48400	13.700	15.599
NORTH	720	14.7056	0.48367	13.757	15.654
SOUTH	719	14.2754	0.48400	13.326	15.225

Std Error uses a pooled estimate of error variance

Figure 5.4: ANOVA tables by scenario variant of Total Failures and failures by type, showing the levels of significance between variants.

Looking at the effect of scenario variants on failures by agent results in similar findings to failures by type. Again, there is a significant difference between variants among the agents that produced higher numbers of failures (Figure 5.6). However, unlike by type, the EAST variant does not always contain the lowest mean. For the CO CDR, the SOUTH variant contains the lowest mean and for the Maneuver PLT Leader, although the EAST variant did contain the lowest mean, the SOUTH variant was insignificantly different. The agents that pull the EAST variant mean down the most are the the UAV and two scout teams. Again, this goes contrary to the distance argument. Since there is no other clear difference between scenario variants, this difference between them may be a result of random variation.

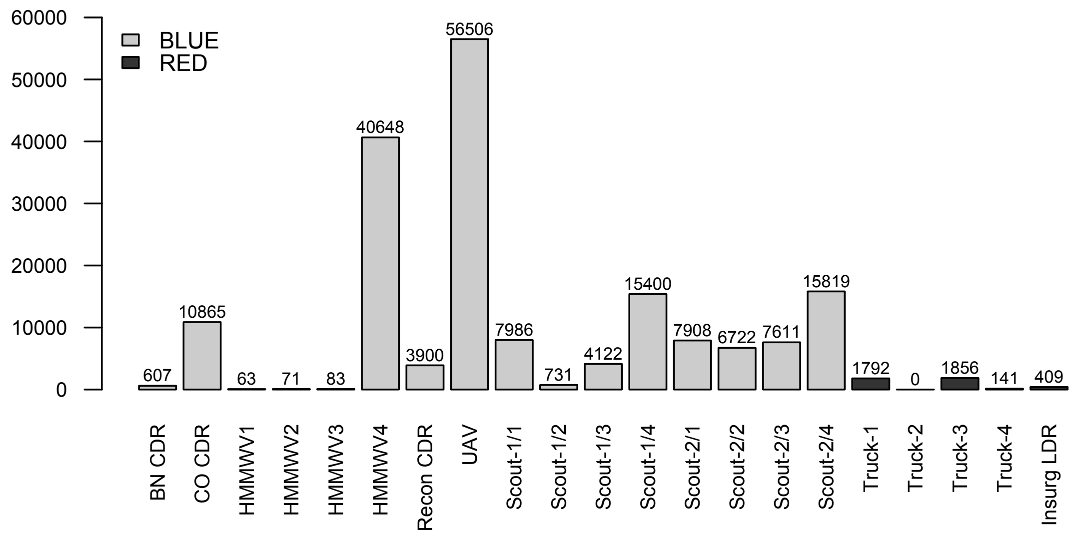


Figure 5.5: Bar chart of Total Failures by agent. This is evidence that an agent's DMP and hierarchical position play a role in the number of failures it encounters.

MANEUVER COMPANY COMMANDER (CO CDR)

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Scenario	2	440.066	220.033	15.2113	<.0001*
Error	2155	31172.328	14.465		
C. Total	2157	31612.393			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
EAST	719	5.53547	0.14184	5.2573	5.8136
NORTH	720	5.12778	0.14174	4.8498	5.4057
SOUTH	719	4.44089	0.14184	4.1627	4.7190

Std Error uses a pooled estimate of error variance

MANEUVER PLATOON LEADER (HMMWV-4)

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Scenario	2	14560.42	7280.21	37.2888	<.0001*
Error	2155	420739.51	195.24		
C. Total	2157	435299.93			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
EAST	719	16.6704	0.52110	15.648	17.692
NORTH	720	22.4861	0.52073	21.465	23.507
SOUTH	719	17.3463	0.52110	16.324	18.368

Std Error uses a pooled estimate of error variance

UAV

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Scenario	2	23174.30	11587.1	67.0790	<.0001*
Error	2155	372252.30	172.7		
C. Total	2157	395426.60			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
EAST	719	21.5522	0.49015	20.591	22.513
NORTH	720	28.6597	0.48981	27.699	29.620
SOUTH	719	28.3380	0.49015	27.377	29.299

Std Error uses a pooled estimate of error variance

SCOUT TEAM 1 LEADER (SCOUT 1/4)

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Scenario	2	194.939	97.4697	5.1921	0.0056*
Error	2155	40455.007	18.7726		
C. Total	2157	40649.946			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
EAST	719	6.71905	0.16158	6.4022	7.0359
NORTH	720	7.27361	0.16147	6.9570	7.5903
SOUTH	719	7.41586	0.16158	7.0990	7.7327

Std Error uses a pooled estimate of error variance

Figure 5.6: ANOVA tables by Scenario Variant of Total Failures by specific agents showing the levels of significance between variants. The variant producing the lowest mean differs between agents.

5.3.2 Failure Influences

One method to identify a parameter's influence is to fit a regression model with communication failures as the response. Using this methodology, two models were built: a simple linear regression model with no interactions and a more complex regression model using all two-way interactions and polynomial terms to degree two. Both models were fit using Total Failures as the response to all of our parameters except *Duration Mode* and *Propagation Mode*, which are both flag parameters for *Bandwidth* and *PCOM*, respectively. Additionally, since it was shown to be significant, scenario variant was also included in the model. To eliminate irrelevant parameters, the complex model was run through a forward stepwise Minimum Bayesian Information Criterion (BIC) process (SAS Institute Inc., 2010). The results of these two models, shown in Figure 5.7 produce very similar results. In both models, *PCOM*, *Max TCP Retries*, and *TCP Wait Time* are the top three parameters contributing to failed communications, with *PCOM* providing the most influence. Scenario variant is also shown to be a significant contributor. In the simple model, *Link Refresh Seconds*, *Antenna Gain*, *Bits in Data Unit*, and *Bits of Overhead* are shown to be insignificant. In the complex model, *Bits in Data Unit* and *Bits of Overhead* remain in the model only for their interactions with other parameters. Furthermore, after the stepwise process, only 20 of the the possible 55 terms are included and *Antenna Gain* is removed altogether. It is also interesting to note, that despite their dependent relationship, the *Bits in Data Unit* and *Bits of Overhead* interaction term is also removed by the stepwise process. However, since much of the response variation is unaccounted for in these models—as evidence of their respective low R^2's of 0.44 and 0.55—the simulation does not appear to be well explained by such regression techniques. That being said, they do help show levels of influence.

It makes sense that *PCOM* is the most significant contributor to failures. A "coin" weighted by the *PCOM* parameter value is essentially flipped to determine if the message will still be successful after the effects of all of the other parameters are applied to a message. If PCOM is too strong of a parameter, what would happen if *PCOM* is taken out of the equation? Subsetting the data to contain only the 44 design points from each variant (660 runs) where *Propagation Mode + PCOM* was "PERFECT + 1.0" helps depict this effect. In the upper left of Figure 5.8, the significance of *PCOM* to Communication Failures is shown by the steepness of the smoother line through the observations. In contrast, the rest of Figure 5.8 shows those same interaction affects with the other parameters when *PCOM* is varied and when it is 1.0. With *PCOM* varied, the steepest smoother lines belong to *Max TCP Retries* and *TCP Wait Time*. However, when

PCOM is 1.0, most of the lines remain fairly flat, with only *Bandwidth* having a strong descent between 0.9 and 1.0. From these graphs, there appears to be a change in influence based on the ranges of the parameters.

SIMPLE REGRESSION MODEL

Term	t Ratio	Prob>|t|
PCOM	-37.97	<.0001*
TCP Retries	-16.90	<.0001*
Scenario[EAST]	-8.87	<.0001*
Scenario[NORTH]	7.68	<.0001*
TCP Wait	6.86	<.0001*
Link Meters	3.70	0.0002*
Bandwidth	-3.63	0.0003*
Link Seconds	1.17	0.2407
Antenna Gain	-1.12	0.2617
Bits Unit	-0.57	0.5658
Bits Overhead	-0.47	0.6374

Summary of Fit	
RSquare	0.471334
RSquare Adj	0.468624
Root Mean Square Error	26.9011
Mean of Response	84.91196
Observations (or Sum Wgts)	2158

COMPLEX REGRESSION MODEL

Term	t Ratio	Prob>|t|
PCOM	-39.28	<.0001*
TCP Retries	-20.56	<.0001*
(PCOM-0.79152)*(TCP Retries-3.24884)	14.00	<.0001*
Scenario{EAST-SOUTH&NORTH}	-9.91	<.0001*
TCP Wait	9.63	<.0001*
(PCOM-0.79152)*(TCP Wait-1.62567)	-8.98	<.0001*
(TCP Retries-3.24884)*(TCP Wait-1.62567)	8.12	<.0001*
(TCP Retries-3.24884)*(TCP Retries-3.24884)	7.97	<.0001*
(Link Meters-23.202)*(Bits Unit-3466.9)	-6.91	<.0001*
(TCP Retries-3.24884)*(Bits Overhead-394.098)	-6.12	<.0001*
(PCOM-0.79152)*(PCOM-0.79152)	5.57	<.0001*
(PCOM-0.79152)*(Bits Overhead-394.098)	-4.75	<.0001*
(TCP Retries-3.24884)*(Bits Unit-3466.9)	4.53	<.0001*
Scenario{SOUTH-NORTH}	-4.15	<.0001*
(PCOM-0.79152)*(Link Seconds-21.2572)	3.70	0.0002*
Bandwidth	-3.56	0.0004*
(TCP Wait-1.62567)*(TCP Wait-1.62567)	3.46	0.0006*
(Scenario{EAST-SOUTH&NORTH}+0.33364)*(PCOM-0.79152)	3.23	0.0013*
(Scenario{SOUTH-NORTH}+0.00046)*(Bandwidth-0.80547)	2.79	0.0053*
(Bandwidth-0.80547)*(PCOM-0.79152)	-2.63	0.0085*
(Scenario{SOUTH-NORTH}+0.00046)*(TCP Wait-1.62567)	-2.18	0.0297*
Link Meters	1.50	0.1348
Link Seconds	-1.02	0.3078
Bits Unit	0.97	0.3299
Bits Overhead	-0.03	0.9745

Summary of Fit	
RSquare	0.577783
RSquare Adj	0.572832
Root Mean Square Error	24.11951
Mean of Response	84.91196
Observations (or Sum Wgts)	2158

Figure 5.7: Simple linear and complex regression models with Total Failures as the response to all parameters except *Duration Mode* and *Propagation Mode*. The sorted parameter estimates show the level of influence each parameter had to the model. The complex model contains all two-way interactions and polynomial terms to degree two remaining after a forward Minimum BIC stepwise process.

A partition tree does a great job of breaking up influence by ranges within each parameter. Again, using Total Failures as the response and all parameters except *Duration Mode* and *Propagation Mode* produces the partition tree shown in Figure 5.9. This tree shows that there is, in fact, a difference in influence depending on the value of *PCOM*. When *PCOM* is varied in the model, *Max TCP Retries* and *TCP Wait Time* produce the greatest effect on the number of failures. However, when *PCOM* is set to 1.0, *Bandwidth* becomes influential when it is equal to 1.0, confirming the graph in Figure 5.8. *Bandwidth* is then followed by *Antenna Gain* and

Bits of Overhead. This tree includes 18 splits and produces an R^2 of 0.6. The tree continues to improve on that value, but begins to flatten out around 0.62, at which point every parameter is included except for *Link Refresh Meters*.

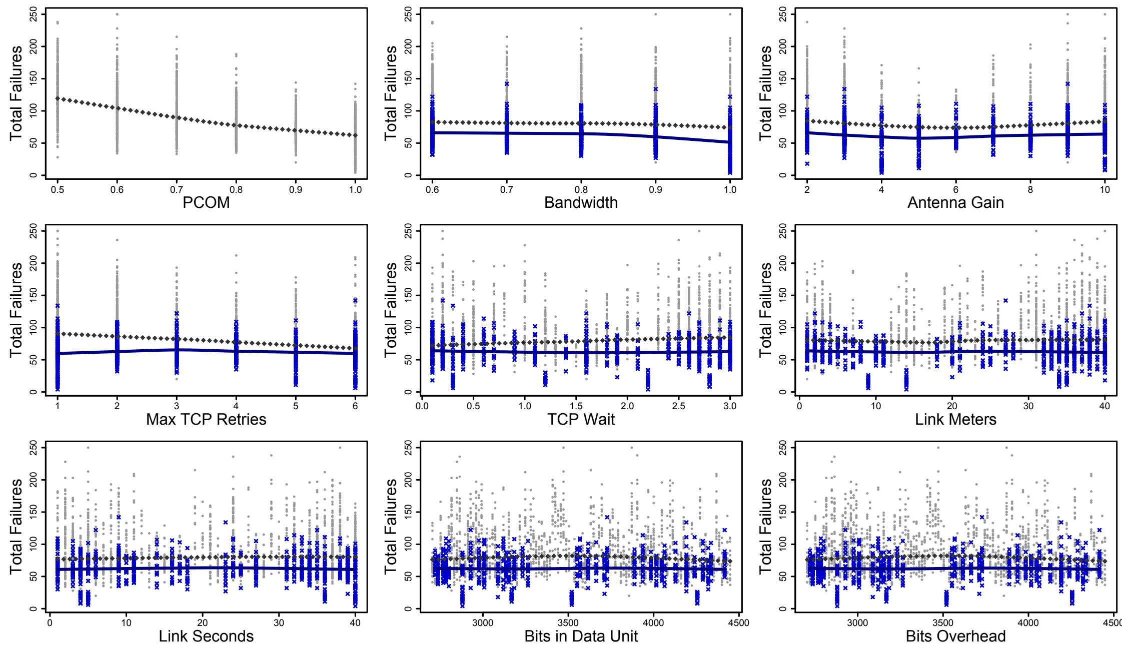


Figure 5.8: Failures by each parameter. The lighter shades and the upper smoother line show the failures with *PCOM* varied in the model, whereas the darker blue shades and the lower smoother line show the failures with *PCOM* equal to 1.0. The slope of the smoother line indicates influence between the parameter and the response. Smoother lines were calculated using a locally weighted scatterplot smoothing (LOESS) smoother to degree 1.

5.4 OPERATIONAL DELAY

Intuitively, the Operational Delay response should have a fairly strong positive correlation with failed communications. That is, as failures increase from a poor communications environment, so should delays in message receipts. Because of this relationship, it seems logical that many of the results seen in the Communications Failure response should be duplicated in Operational Delay. However, an estimate of the correlation ($\hat{\rho}$) when the Operational Delay could be calculated is approximately 0.264. A quick look at the two responses plotted together further suggests that this correlation is not strong. The plot, shown in Figure 5.10, confirms this is a positive correlation with a smoother line through the data that indicates an increasing trend, but the increase is very slight and surrounded by a lot of scatter and variability. In addition to revealing instances where the change of mission was never received (represented by the *"x's"* at the top of the plot), the plot also shows multiple occasions where the calculated delay is negative, indicating a possible error with the Operational Delay formula. It may be the case that the behavior of Operational Delay is more related to its calculated values made from a variety of assumptions pulled from various outputs, rather than a direct reflection of the simulation.

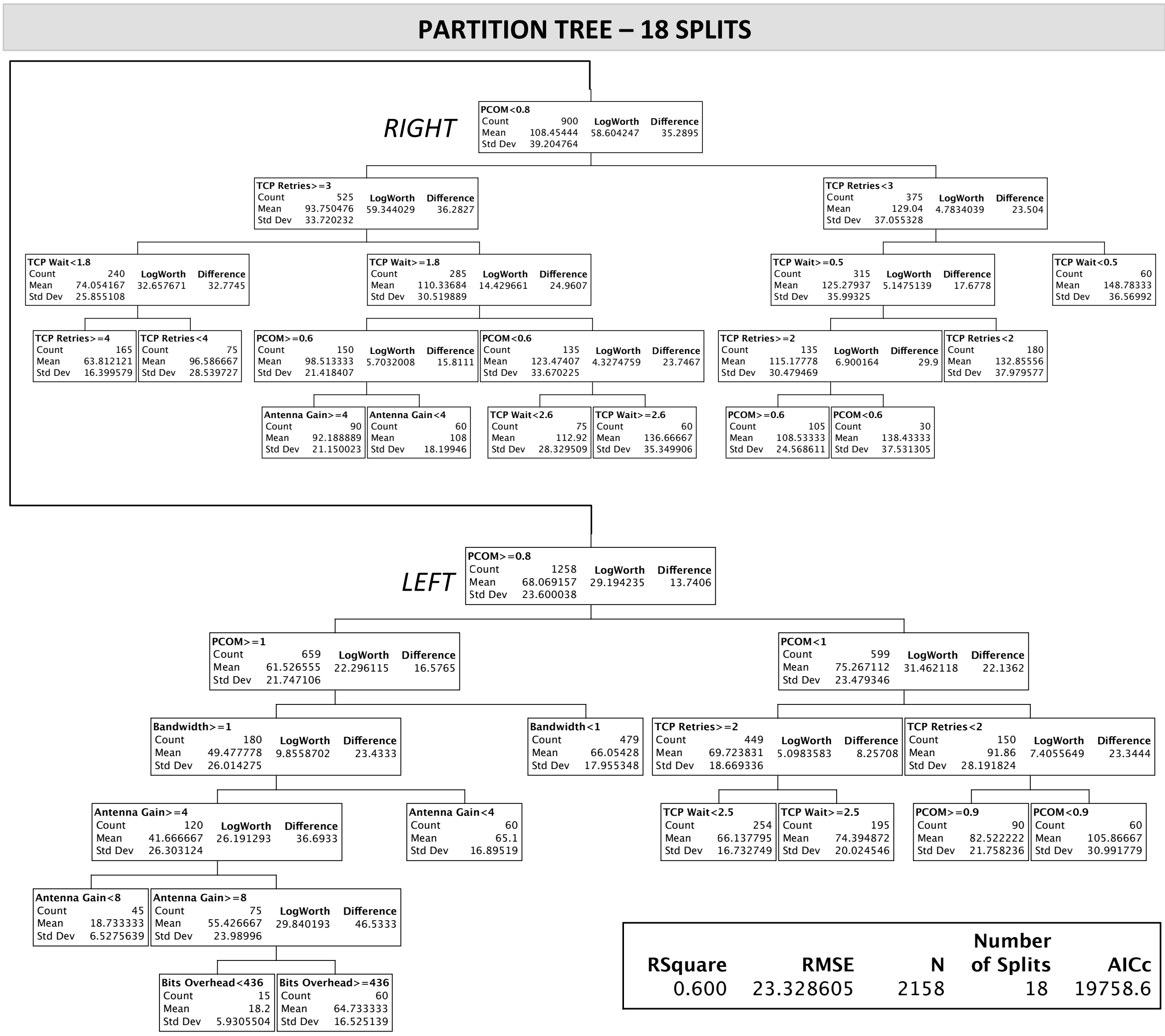


Figure 5.9: Partition tree of Total Failures as the response to all parameters except *Duration Mode* and *Propagation Mode*. The tree shows that different parameters have influence on the model, depending on the value of *PCOM*. In each box, "count" refers to the number of instances and "mean" refers to Total Failures for a given partition.

5.4.1 Change of Mission not Received

The experiment produced 82 instances in which the change of mission MIP was not received. Possible explanations for these instances include: (1) the order was issued by the BN CDR, but not received by the PLT Leader; (2) the BN CDR was not informed of RED's presence in the

AOI; or (3) the BN CDR received the information too late and decided to let the Maneuver PLT continue to its initial OBJ due to its close proximity. Although the output and tools available for this thesis prevent tracking down the actual explanation of each instance, there is enough data to make some logical inferences. From the output, there were 11 and 10 instances when Insurgent Teams 1 and 2 never initiated movement, respectively. However, since there were no instances when both insurgent teams failed to start, the instances where the MIP was not received must be a result of communications failures within the BLUE networks. A comparison of failures of BLUE agents by whether or not the mission change was received (Figure 5.11) reveals a dramatic difference. Each agent encountered more communication failures during the runs in which the MIP was not received. An analysis of the means of Total Failures in an ANOVA table (Figure 5.12) reveals that this difference is significant. If the communications environment during these instances was, in fact, poor, then the parameter values for these 82 instances should follow our earlier analysis of communications failures.

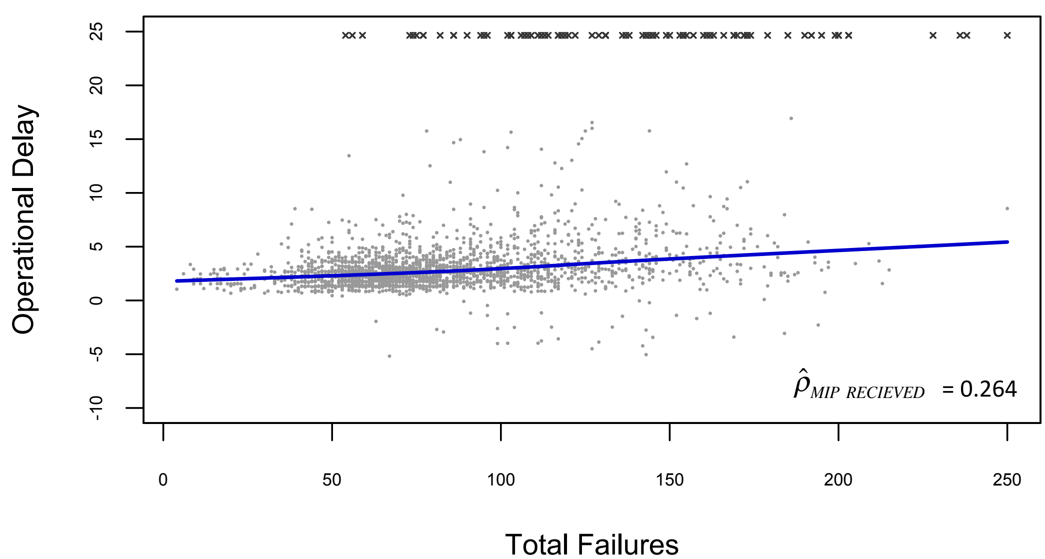


Figure 5.10: Total Failures by Operational Delay. The “x’s” at the top refer to the 82 instances when Operational Delay could not be calculated due to the MIP not received. For the remaining instances, the estimated correlation between the two responses is 0.264. The smoother line is calculated using LOESS to degree 1.

The mean number of Total Failures for these instances is 139.65, with a 95% Confidence Interval of the mean of [130.02, 149.28]. Plotting our two most influential parameters together, *PCOM* and *Max TCP Retries* in Figure 5.13, reveals the makings of a poor communications environment. *PCOM* was mostly at its lowest level of 0.5 and never exceeded 0.8. Similarly, *Max TCP Retries* was predominately at 1.0, with one instance above 3 at 5 retries. The largest combination of the two parameters is when *PCOM* is 0.5 and *Max TCP Retries* is 1. Looking at boxplots of the values of the other parameters for these 82 instances appears to further confirm their diminished importance. Although each tends to lean toward its more restrictive values (except for *Link Refresh-Meters* and *Bits in Data Unit*), their values also span their respective ranges. Regardless, it is clear that communications failures had an impact on whether or not the MIP was received.

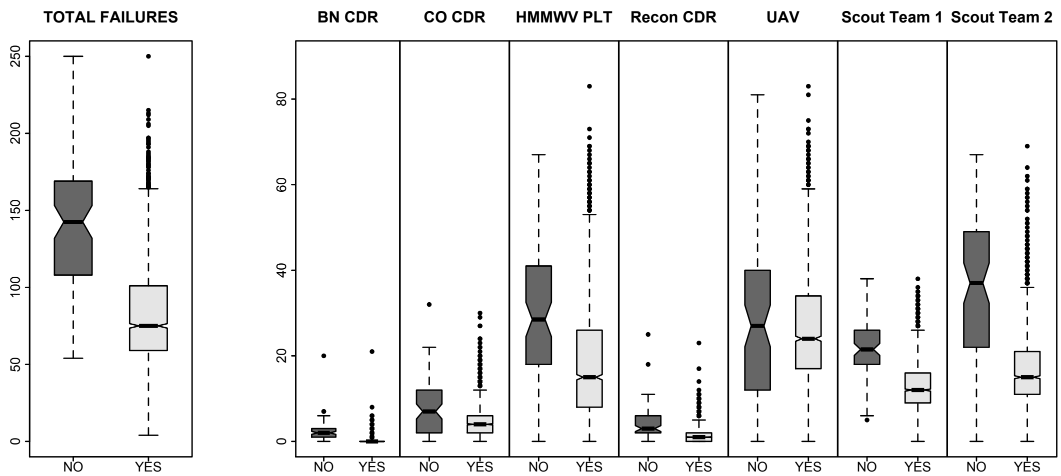


Figure 5.11: Difference in communications failures when the MIP was and was not received. "No" and the darker shaded boxes refers to instances when the MIP was not received.

TOTAL FAILURES

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
CM Received	1	255363.3	255363	205.2651	<.0001*
Error	2156	2682206.0	1244		
C. Total	2157	2937569.3			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
NO	82	139.646	3.8951	132.01	147.28
YES	2076	82.750	0.7741	81.23	84.27

Std Error uses a pooled estimate of error variance

Figure 5.12: ANOVA table comparing the difference between Total Failures when the MIP was and was not received. The F statistic shows that the difference between the two means is significant.

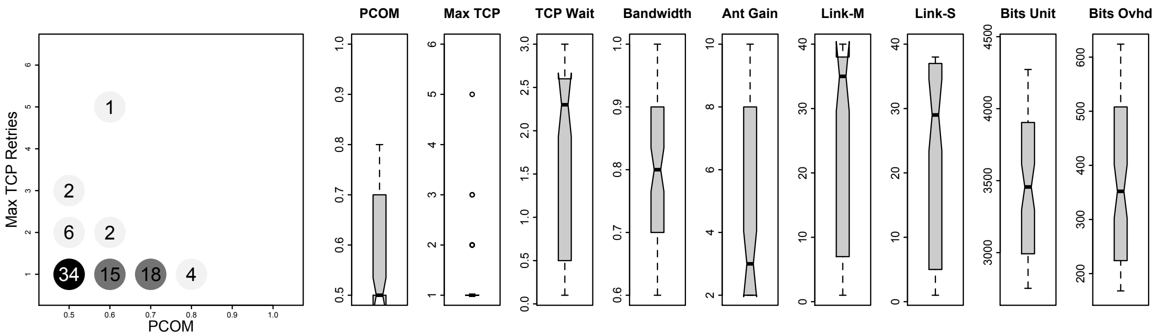


Figure 5.13: Parameter values when the change of mission was not received. Far left, an interaction plot between *PCOM* and *Max TCP Retries* showing that the paired values when the MIP was not received reflect a "poor" communications environment.

5.4.2 Negative Operational Delay

A histogram of the Operational Delays—leaving out the 82 instances when the Operational Delay could not be calculated (shown in Figure 5.14)—confirms the presence of negative calculated delay times. There exists the possibility that the assumption of consistent travel time may not have been accurate and that RED moved faster in some scenarios than in others. Another possibility is that the BN CDR issued the MIP before RED entered the AOI, based on reported locations, headings, and speeds of the the RED agents. Based on the modeled DMPs and stochastic nature of the model, both instances are possible. Recall that while both the RED start times and BLUE MIP receipt times were extracted from the data, the travel time was estimated based on the mean travel time in the 12 Seed Variance runs. Since the travel time for each run was not extracted, a comparison between our estimate and the mean travel time for our experiments cannot be made. However, it may be possible to interpret a comparison based on the statistic results of the two RED Insurgent Team start times and the BLUE MIP receipt times from the Seed Variance runs.

Under a "good" communications environment, the Seed Variance runs produced a mean delay time of 1.83 minutes, with a 95% Confidence Interval of the mean of [1.55, 2.11]. The mean delay time from our experiment with varying communications environment qualities is 2.97 minutes. Although this value is outside our estimated confidence interval, due to the different environments produced by the DOE, it stands to reason that our estimated mean is slightly better than the experiment mean and does not necessarily indicate a formulation error.

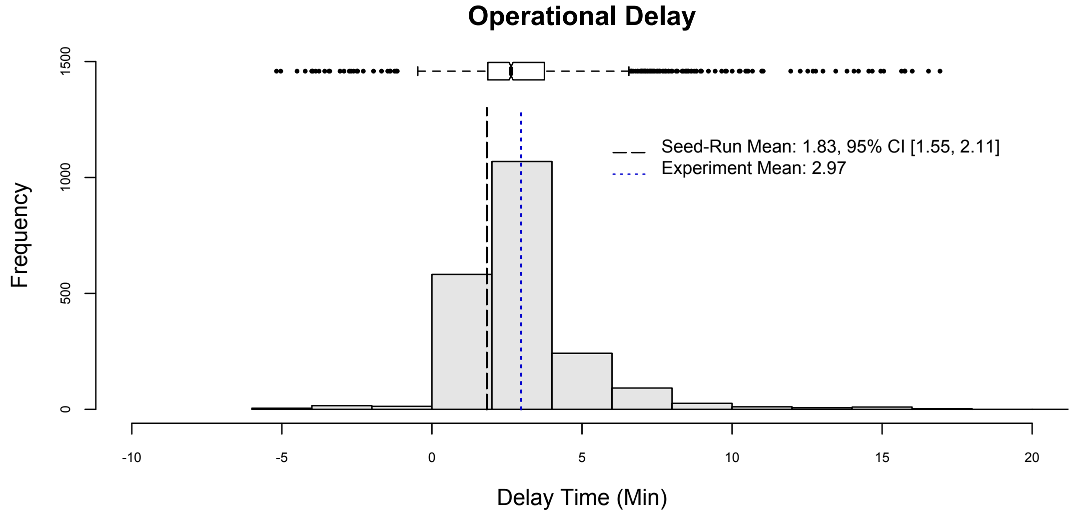


Figure 5.14: Histogram of Operational Delay with corresponding boxplot. The dashed vertical line shows the location of the mean produced by the Seed Variance runs and the blue dotted vertical line shows the location of the experiment mean.

Comparing the start times for both RED teams as well as BLUE MIP receipt times between the Seed Variance runs and the experiment runs produces similar results. The experiment means are to the right and outside of the confidence intervals of the calculated means from the Seed Variance runs. But again, this makes sense based on the changing communications environment. The difference in the distributions between the two start times and the MIP receipt times, however, is interesting. Shown in Figure 5.15, there are no instances in which RED started earlier than their modal value. In fact, each RED team began movement a mere 61 times earlier than the lower confidence interval from the Seed Variance mean estimate. On the other hand, the BLUE MIP was received earlier than the modal value on numerous occasions and 526 times sooner than the lower confidence interval of the Seed Variance mean estimate. Again, this makes sense because these events occur at different times and under different circumstances in the simulation. The RED teams are mostly operating in a similar situation each time: early in the simulation, agents are close together, and there is little traffic to potentially clog the communications networks. The BLUE change of mission occurs much later in the simulation, with different distances between agents, and plenty of messages to potentially clog the networks. However, the distribution of the change of mission time more closely resembles the overall Operational Delay distribution. Understanding the environment that produced earlier changes to the mission might explain the negative Operational Delays.

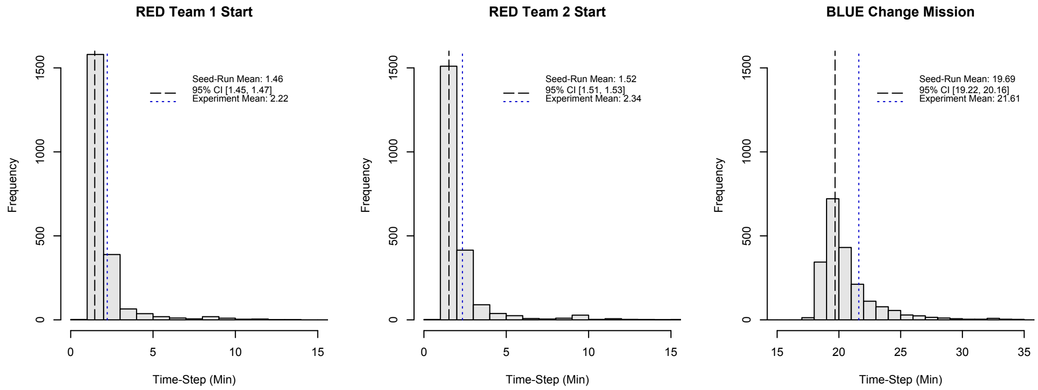


Figure 5.15: Histograms of RED Start Times and BLUE MIP Times. The dashed vertical lines show the location of the means produced by the Seed Variance runs and the blue dotted vertical lines show the location of the experiment means.

The expected cause should be the opposite of what we saw in the instances when the MIP was never received. That is, the communications environment should be great. In fact, it should be better than the environment in the Seed Variance runs, which produced a mean Total Failures of 14.58, with a 95% Confidence Interval of the mean of [12.63, 16.53]. Figure 5.16 replicates Figure 5.13 by using the 856 instances to the left of the Seed Variance mean MIP receipt time of 19.69. From the interaction plot between *PCOM* and *Max TCP Retries*, although we see instances where the values were worse than the Seed Variance "base" values (Table 4.1), the majority of the runs occur when *PCOM* is equal to 1.0, and the majority of those occur when *Max TCP Retries* meets or exceeds the base value of 3. Furthermore, the boxplot of *TCP Wait Time* leans above its base value of 1.0, which was shown to improve the communications environment when *PCOM* was less than 1.0. However, these instances produced a mean Total Failures of 68.32, with a 95% Confidence Interval of the mean of [66.49, 70.15]. Although this value is low in consideration of the overall distribution of Total Failures, it is much higher than the mean from the Seed Variance runs. From this evidence, it appears as though the calculated negative values are indeed a product of a very good communications environment, but there also appears to be some error caused by the constructed formula.

5.4.3 Delay Influences

It was argued earlier that there should not only exist a positive correlation between Communications Failures and Operational Delay, but that as a result, their influences should be the same. This hypothesis is easily disproved by duplicating the models produced for Communications Failures. Although *PCOM* remains the top influence, *Max TCP Retries* is now almost as influential. The remaining parameters and/or interaction terms all find different levels of influence. Even some parameters that were significant in predicting Communication Failures are now insignificant when predicting Operational Delay. For example, *Antenna Gain* was left out of the complex model for Communication Failures, but is not removed by the stepwise process in the complex model for Operational Delay. Figure 5.17 shows a lot of straight smoother lines through the output by each parameter. Only *PCOM* has any noticeable slope. A partition tree of the data reveals a similar result. Shown in Figure 5.18, after one split by *Max TCP Retries*, the remaining six splits all involve *PCOM*, suggesting a highly non-linear relationship between *PCOM* and Operational Delay. However, in this tree—as in the two regression models—R^2 is extremely low. Although part of this low value can be attributed to the handling of instances when Operational Delay could not be calculated, the great amount of variance is evident in Figure 5.10 with or without those 82 instances.

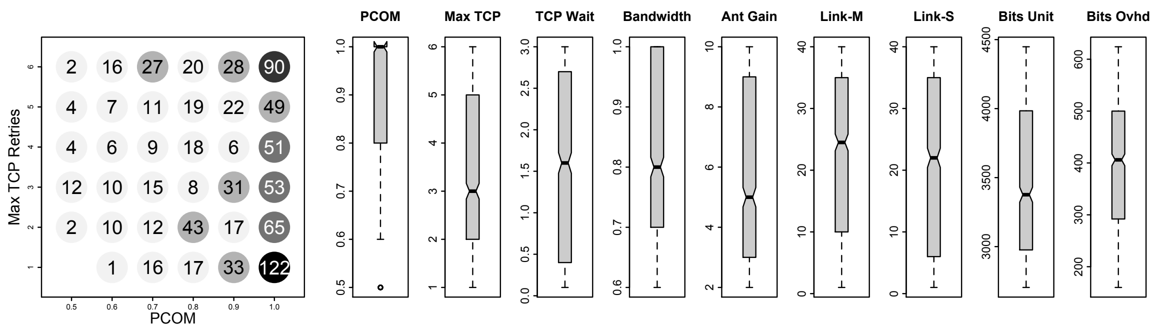


Figure 5.16: Parameter values when the MIP was received earlier than the Seed Variance mean time. Far left, an interaction plot between *PCOM* and *Max TCP Retries* showing that the paired values, when the MIP was not received, reflect a "good" communications environment.

Consider the experimental design: the communication parameters were varied for *all* agents, both BLUE and RED. Therefore, Operational Delay is calculated using two values affected by the communications environment. When the communications environment results in BLUE taking longer to receive the MIP change, that same environment no doubt resulted in RED taking longer to initiate movement. The difference between the two times may still change as a result of the changing environment, but since they are moving together in the same direction the impact is less discernible. This realization, in addition to the variance and noise from the travel time calculation, means that parameter influences in regard to Operational Delay cannot be evaluated in too much detail. In hindsight, it would have been beneficial to only vary BLUE's parameters and keep RED's constant. That being said, there is enough evidence to support the influence of both PCOM and Max TCP Retries on Operational Delay.

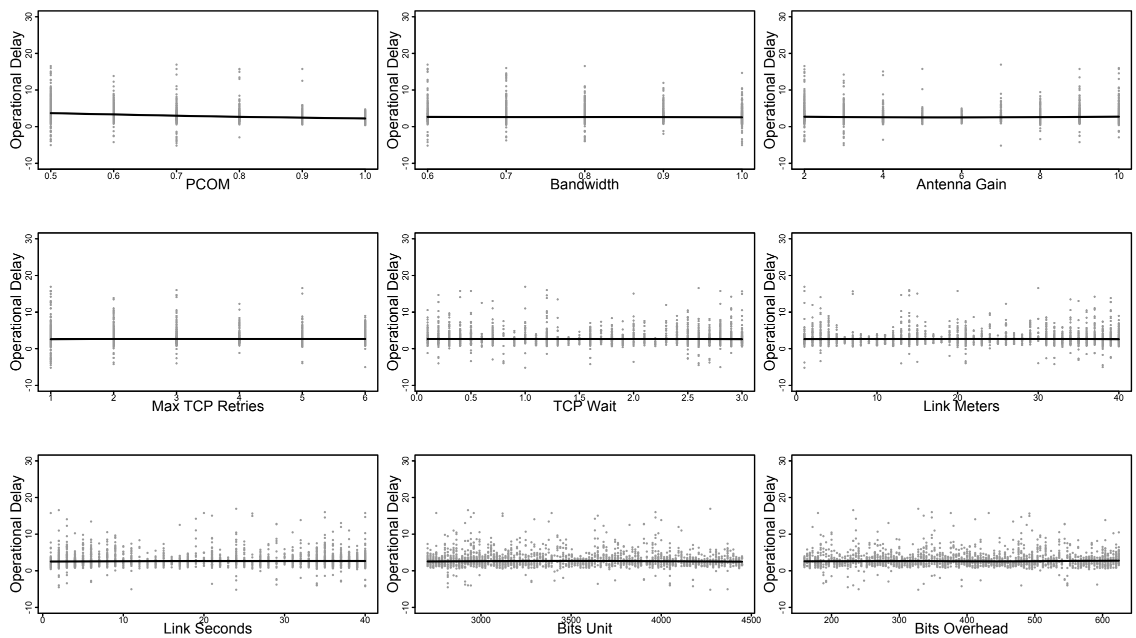
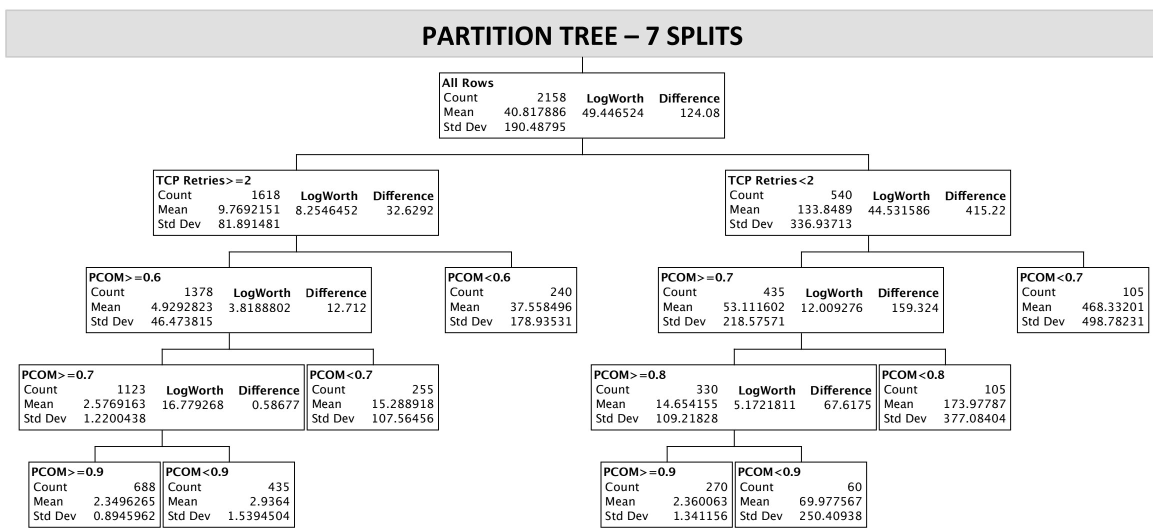


Figure 5.17: Operational Delay by each parameter. The slope of the smoother line indicates influence between the parameter and the response. Smoother lines were calculated using a LOESS smoother to degree 1.



RSquare	RMSE	N	Number of Splits	AICc
0.298	159.5787	2158	7	28035.3

Figure 5.18: Partition tree with Operational Delay as the response to all parameters except *Duration Mode* and *Propagation Mode*. In each box, "count" refers to the number of instances and "mean" refers to Total Failures for a given partition. The tree shows that *PCOM* has the majority of influence on the response; however, the tree does produce a low R^2.

CHAPTER 6: CONCLUSIONS AND RECOMMENDATIONS

This thesis set out to explain key communications factors within S4 with quantitative analysis. By using a focused scenario, the data farming process, incorporating a cutting edge DOE, and various analytical models and tools, this thesis answered some critical questions regarding the communication environment within the simulation and paved the way for future studies and a more in-depth VV&A process. The conclusions reached, however, are limited to the scenario examined over the parameter ranges explored.

6.1 RESEARCH QUESTIONS

In developing the purpose of this thesis, three primary questions were formulated to address the concerns raised by NMSU-PSL regarding S4:

1. Which communications factors have the most influence on the model's output?
2. How sensitive are the model's DMPs to the success or failure of communications?
3. Is the model's output a result of the equipment capabilities being analyzed or a product of how the communications environment is modeled in S4?

The following sections discuss the results of our analysis in consideration of these questions.

6.1.1 Influence

Two responses and a variety of techniques, including regression models and partition trees, were used to show how parameters influenced the model. For the most part, results varied between techniques and responses. Some models threw out parameters and/or interaction terms that others retained. However, each technique and response had one thing in common: they each listed *PCOM* and *Max TCP Retries* as the most influential values in determining the response value for this scenario. Additionally, looking at the one-way effects, those two parameters had influence across their entire varied range, whereas the impact of others existed with less regard to changes in their value. This does not necessarily imply that the other parameters are useless in the model. These influence techniques left a lot of unexplained variance shown by their low R^2 values. Additionally, partition trees provided evidence that different parameters have different levels of influence based on the overall communications environment. For instance, *Bandwidth* plays a greater role when *PCOM* is not equal to 1.0. However, this conclusion points to two possibilities worth exploring: (1) modeling of some parameters may not be consistent with real-world effects or (2) some parameters may not be necessary—at least in the scenario modeled.

There were no adverse interactions identified; however, the wide disparity between influence leads one to wonder if maybe some parameters are modeled too strongly and/or others too weakly. Is the *PCOM* "coin flip" located in the proper place within the model? Does *Link Refresh* have a direct, real-world impact on communications success or are its effects truly minimal and only noticed in its interaction with other factors? The lesser influence could also be a result of extreme ranges from this stress test. In other words, the effect of *Bandwidth* might be higher if the experiment explored the parameter more closely between 0.9 and 1.0. On the other hand, if the parameters correspond accurately to real-world effects, yet their impact is so low that the resulting output is barely affected by it presence, it may not be worth the additional computing power to include them in the simulation. A simpler model that produces the same results and allows for the same analysis as a complex model can save tremendous time and effort, and possibly provide the user with faster and/or more robust analysis.

6.1.2 Decision-Making Processes (DMP) Sensitivity

This thesis did not vary any of the decision templates. Instead, it attempted to determine how sensitive the modeled DMPs are to failed communications by analyzing a quantifiable delay caused by poor communications. If the simulation produced operational results without respect to communications, then the entire simulation fails as a network-centric model. Unfortunately, a flaw in the experiment design limited the analytic power of the created Operational Delay response, thus preventing an actual calculation of sensitivity. However, through all of the noise there exists significant evidence showing that the quality of the communications environment did, in fact, play a role in the timing of the DMPs. In a "good" communications environment, the change of mission order could be received up to 6 minutes earlier than the base case and as late as 15 minutes in a "poor," but still effective environment. When the environment was completely ineffective, the key decision to redirect the Maneuver PLT was either not made or not received as a result of a "bad" communications environment. Based on these results, it is clear that decisions in S4 cannot be made without communications and that the quality of the communications environment affects the outcome of the simulated operation.

6.1.3 Equipment Versus Environment

The output from this thesis reflects only the effects of key communications environment parameters and did not extract any response to directly analyze the modeled equipment effects. However, similar to the rationale in drawing a conclusion regarding the DMP, if the communications environment was ineffective in improving or degrading communications, it can be argued that the modeled equipment is unrealistically operating without restriction. Of course, the results showed a wide range of failures stemming from the changing communications environment. It is apparent that the communications environment is indeed playing a critical role between equipment sending and receiving communications. However, since communications equipment was not varied, this conclusion is limited only to the equipment modeled in this experiment.

6.2 RECOMMENDATIONS FOR FUTURE STUDY

This thesis just began to scratch the S4 surface. The simulation contains thousands of parameters belonging to multiple processes. In addition to communications, the developers at PSL would benefit greatly from detailed analysis regarding the sensing, terrain, movement, decision making, and ballistics engagements within S4. Since this thesis used only a handful of parameters, further analysis regarding the communications environment is also needed. Furthermore, an additional study is needed to explore these same parameters under a different set of less stressful ranges that might better emphasize or explain interactions. The unexplained difference between scenario variants also warrants further exploration to ensure the effect of distance is accurately portrayed in the model. Most importantly, the methodologies outlined in this thesis can be used as a guide for future research of either another selected group of parameters or all of the parameters belonging to a particular process. Future studies, however, should take into consideration the challenges this thesis faced regarding the calculation of Operational Delay, random number generation, lack of cluster computing, and output handling.

6.2.1 Multiple Replications and Cluster Computing

Future studies must incorporate multiple replications of S4 due to the substantial variability found across design points and within replications. Doing so, in conjunction with a DOE, will ensure robust answers over a breadth of possibilities. In order to accomplish this, cluster computing and a sophisticated DOE, should be used. The lack of access to a computer cluster to conduct the runs for this thesis limited the total number of runs that could be conducted. A cluster computer would have enabled either a larger and more space-filling DOE and/or more replications per design point, thereby improving the accuracy of the results. Additionally, it would have enabled this thesis to follow the "iterative" process of data farming. Instead, the amount of time needed to conduct the runs removed any possibility for reruns (possibly fixing the "RED Constant" issue) or the opportunity to "zoom" in on some of the more interesting interactions or ranges. Future studies that intend to analyze more parameters at once should strongly consider using a computer cluster to ensure that enough design points and replications are used.

6.2.2 Random Number Generation

The developers at PSL are working on understanding when and in what order each process draws from the random number stream. Although this would help to explain the uncertainty of the simulation's stochastic processes, it might prove more beneficial to use a separate stream of random numbers for each process within the simulation. Doing so would make future studies such as this, as well as the VV&A process, much simpler. With a separate streams of numbers, independence between runs can be guaranteed and variance reduction techniques can be applied in future studies.

6.2.3 Leave RED Constant

The difficulty regarding the calculation and analysis of the Operational Delay response could have been easily mitigated. Had the experimental design only varied the communications parameters for the BLUE force and left the RED force constant, there would have been far less variability in RED's communications and movements. This would simplify the formulation for Operational Delay since RED would initiate movement and arrive in the AOI within a tighter time interval. Additionally, the noise in the calculation would be diminished, since RED's changing movement would no longer have correlation to BLUE's MIP receipt.

6.2.4 Output Handling

The fact that the runs for this thesis were conducted dislocated from the researcher limited the amount of output available for analysis. Due to the size of the output, not all of the files could be reasonably transported between PSL and the researcher. Furthermore, access to the advanced postprocessing tools of PSL was also limited. Access to all of the output files and PSL's postprocessing tools would increase the accuracy and variety of response parameters and improve the detail of future analysis.

6.3 RESEARCH SUMMARY

This thesis sought to explore the interactions of selected communications parameters within S4. It does not purport to possess or provide expertise regarding real-world communications effects. The varying levels of influence found may be a legitimate representation of reality and the effects of some of the lesser influential parameters may provide a greater impact under a different scenario or sets of ranges. What is most important is that S4 reacted in a logical manner—as the communications environment degraded, so did the ability of agents to communicate with each other—and that ability improved as the communications environment improved. Such a realization of performance cannot be said of all simulations during development. S4 is clearly heading in the right direction. Its continued exploration and development incorporating the analysis, methodology, and lessons learned from this thesis should pave the path toward VV&A.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A:
NOLH DESIGN MATRICES

The following matrices belong to the 33- and 256-design point NOLH designs that were considered—but not used—for this thesis. The disparity between these two designs and the NONBMD that was used, show the importance of using sophisticated DOE for future S4 research.

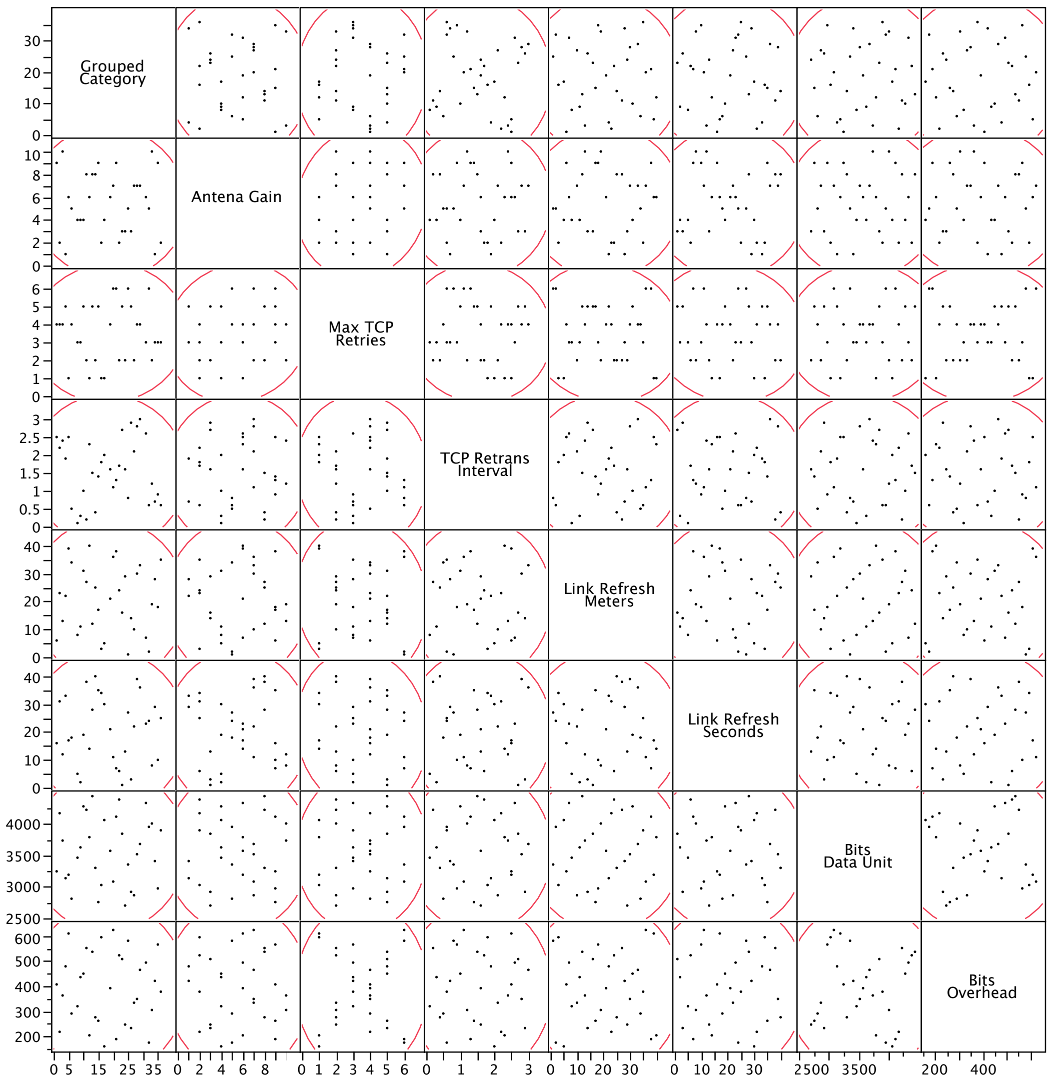


Figure A.1: 33-design point NOLH provides less design space coverage, produces higher ρ_{map} than the NONBMD, and does not incorporate the bit restriction rule. Note the corner lines in each of the design space blocks depict the reduced number of samples near the corners.

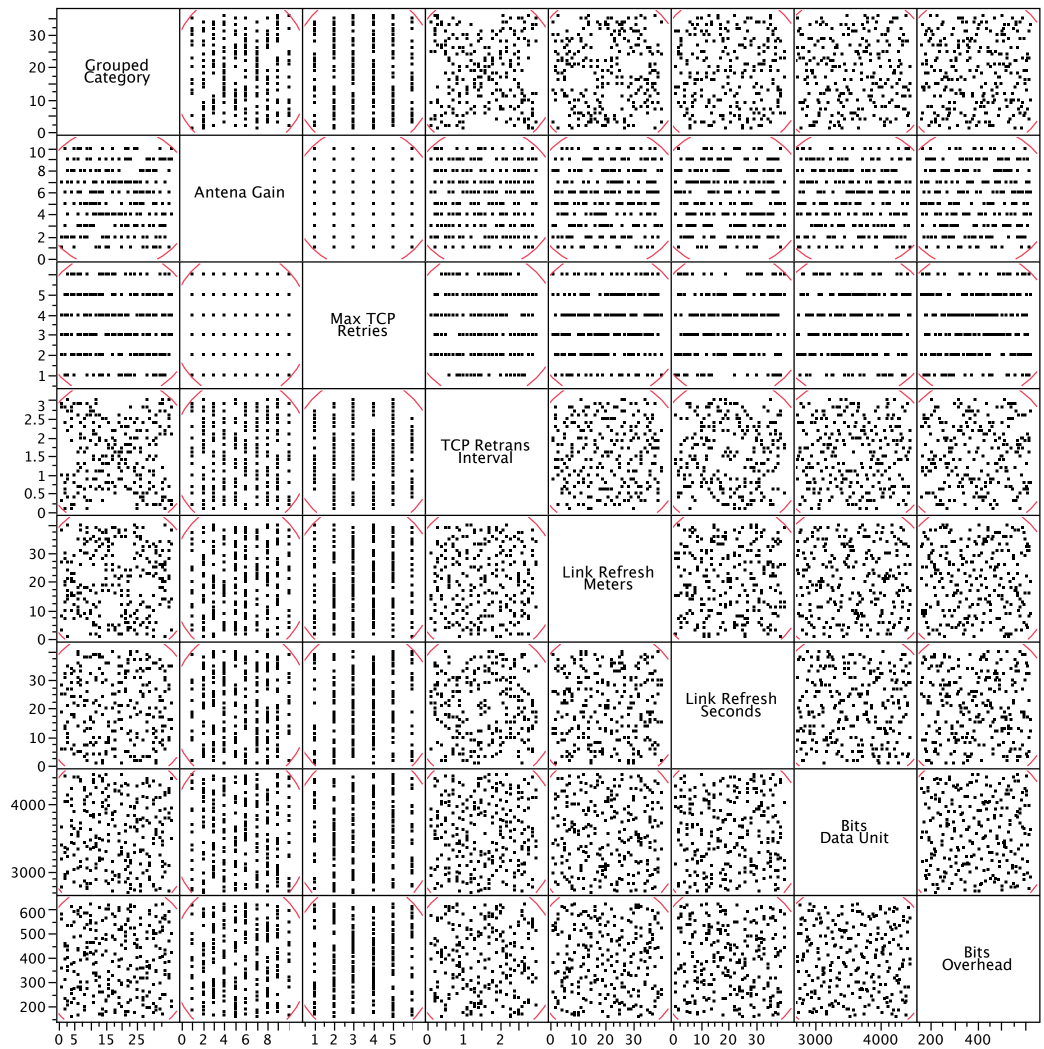


Figure A.2: 256-design point NOLH provides slightly less design space coverage, produces a slightly higher ρ_{map} than the NONBMD, and does not incorporate the bit restriction rule. Note the corner lines in each of the design space blocks depict the reduced number of samples near the corners.

LIST OF REFERENCES

Army Transformation Office (2004). *Army transformation roadmap*. Office of the Deputy Chief of Staff, U.S. Army Operations, Washington, D.C.

Bernstein Jr., R., Flores, R., & Starks, M. (2006). *Objectives and capabilities of the system of systems survivability simulation (S4)*. Technical Report ARL-TN-260, Army Research Laboratory.

Cioppa, T. M. & Lucas, T. W. (2007). Efficient nearly orthogonal and space-filling Latin hypercubes. *Technometrics*, *49*(1), 45–55.

Committee on Modeling & Simulation for Defense Transformation (2006). *Defense Modeling, Simulation, and Analysis: Meeting the Challenge*. Washington, D.C.: The National Academies Press.

Davidson, J. & Pogel, A. (2010). Tactical agent model requirements for M&S-based IT: C2 assessments. *C2 Journal*, *4*(1), 1–52.

Davidson, J., Pogel, A., & Smith, J. A. (2008). The role of battle command in information system assessments. In *Proceedings from the International Conference on Industrial Engineering Theory, Applications, and Practice*, (pp. 154–160).

Department of Defense (2009). *DoD modeling and simulations (M&S) verification, validation, and accreditation (VV&A)*. Technical Report DoDI 5000.61, Under Secretary of Defense for Acquisition, Technology, and Logistics.

Hartley, R. (2010). Decision making in the system of systems survivability simulation. NMSU-PSL Working paper.

Horne, G. E. & Meyer, T. E. (2004). Data farming: Discovering surprise. In *Proceedings from the Winter Simulation Conference*, (pp. 807–813).

Kleijnen, J. P. C., Sanchez, S. M., Lucas, T. W., & Cioppa, T. M. (2005). A user's guide to the brave new world of designing simulation experiments. *INFORMS Journal on Computing*, *17*(3), 263–289.

Law, A. M. & Kelton, W. D. (2000). *Simulation, modeling, and analysis* (3rd ed.). Boston, Massachusetts: McGraw-Hill Higher Education.

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2006). *Introduction to linear regression analysis* (4th ed.). Hoboken, New Jersey: Wiley-Interscience.

Sanchez, S. M. (2008). Better than a petaflop: The power of efficient experimental design. In *Proceedings from the Winter Simulation Conference*, (pp. 73–84).

SAS Institute Inc. (2010). *JMP 9 modeling and multivariate methods*. Cary, NC: SAS Institute Inc.

Smith, J. A., Bernstein Jr., R., Hartley, R., & Harikumar, J. (2010). System of systems survivability simulation (S4). SLAD and PSL Presentation.

Starks, M. & Flores, R. (2004). *New foundations for survivability lethality vulnerability analysis (SLVA)*. Technical Report ARL-TN-216, Army Research Laboratory.

Vieira Jr., H. & Sanchez, S. M. (2010). Efficient nearly orthogonal nearly balanced mixed designs. Working paper.

Vong, T. T., Haas, G. A., & Henry, C. L. (1999). *NATO refence mobility model (NRMM) modeling of the DEMO III experiment unmanned, ground vehicle (XUV)*. Technical Report ARL-MR-435, Army Research Laboratory.

Ye, K. G. (1998). Orthogonal column Latin hypercubes and their application in computer experiments. *Journal of the American Statistical Association*, *93*, 1430–1439.

Zacharias, G., MacMillan, J., & Hemel, S. B. V. (2008). *Behavioral Modeling and Simulation: From Individuals to Societies*. Washington, D.C.: National Academies Press.

Initial Distribution List

1. Defense Technical Information Center
   Ft. Belvoir, Virginia
2. Dudly Knox Library
   Naval Postgraduate School
   Monterey, California
3. National Research Council
   Washington, D.C.
4. Army Research Laboratory
   Adelphi, Maryland
5. Survivability/Lethality Analysis Directorate, Army Research Laboratory
   White Sands Missile Range, New Mexico
6. Physical Science Laboratory, New Mexico State University
   Las Cruces, New Mexico